2010

Profiling and Consent: Stops, Searches and Seizures after *Soto*

Jeffrey Fagan
Columbia Law School, jfagan@law.columbia.edu

Amanda Geller
abg2108@columbia.edu

Workshop Draft
Please do not quote or cite

Profiling and Consent:
Stops, Searches and Seizures after *Soto*[*]

Jeffrey Fagan
Columbia Law School

Amanda Geller, Ph.D.
Columbia University School of Social Work

November 2010

Please address all correspondence to:

Jeffrey Fagan
Columbia Law School
435 West 116th Street
New York, NY 10027
Tel. 212.854.2624
Fax 212.854.7946
Email Jeffrey.Fagan@law.columbia.edu

[*] This research was supported by Grant No. 16395 from The Atlantic Philanthropies, and institutional support from Columbia Law School. We are grateful to the *New Jersey Advisory Committee on Police Standards,* the Office of the Attorney General of the State of New Jersey, and the New Jersey State Police for providing the data for this research. Thanks to Rick Brooks, Chris Winship and Geoff Alpert for advice in the early stages of this project. Special thanks to Andrew Case for excellent research assistance. All opinions, conclusions and any errors are the responsibility of the authors.

Abstract

Following *Soto v State* (1999), New Jersey was among the first states to enter into a comprehensive Consent Decree with the U.S. Department of Justice to end racially selective enforcement on the state's highways. The Consent Decree led to extensive reforms in the training and supervision of state police troopers, and the design of information technology to monitor the activities of the State Police. Compliance was assessed in part on the State's progress toward the elimination of racial disparities in the patterns of highway stops and searches. We assess compliance by analyzing data on 257,000 vehicle stops on the New Jersey Turnpike by the state police from 2005-2007, the final months of the Consent Decree. Specifically, we exploit heterogeneity of officer and driver race to identify disparities in the probability that stops lead to a search. We assume a crime-minimizing or welfarist rationale for stops, under which race-neutral factors are equally likely to motivate stops, regardless of driver or passenger race. We also test a Fairness Presumption (Durlauf, 2006) by comparing search patterns between driver-officer pairs where the driver and officer are different races, and a set of race-neutral benchmarks where the driver and officer are the same race.

Results of fixed effects logistic regressions show that Black and Hispanic drivers, when stopped, are more than twice as likely as White drivers to be searched, regardless of officer race. The results also suggest that search patterns vary significantly by officer race: Black officers are less likely to conduct a search in the course of a stop than are white drivers. We also see significant interactions between the race of officers and that of the drivers they stop: Black drivers are significantly more likely to be searched by white officers than they are by black officers; on the other hand, Hispanic drivers are significantly less likely to be searched by either black or white officers than they are by Hispanic officers. Racial disparities in the selection of stopped drivers for search suggest that despite institutional gains under the Consent Decree in management and professionalization, there were no tangible gains in distributional equity. We review the design of the Consent Decree and the accompanying oversight mechanisms to identify structural weaknesses in external monitoring and institutional design in the oversight of the State Police that compromised the pursuit of equality goals.

I. Introduction

Following *Soto v State* (1999)[1], the U.S. Department of Justice and the State of New Jersey entered into a consent decree to settle complaints filed by the DOJ under 42 U.S.C. § 14141 and 42 U.S.C. § 3789d(c), alleging claims of racial profiling by the troopers of the New Jersey State Police (hereafter, NJSP).[2] The Consent Decree requires that the troopers not rely on the race or national origin of drivers except when they have been given a particular description of a driver that included race or national origin, and required documentation of stops and managerial review to ensure compliance.[3]

New Jersey was among the first states to enter into a comprehensive and far-reaching Consent Decree with the U.S. Department of Justice to end racially selective enforcement on the state's highways. The legal claim against the State and the NJSP came at a time when public and political awareness of racial profiling had saturated public policy debates on race and policing.[4] Consent decrees has been signed in a handful of municipalities, often to remedy problems of excessive use of force and weaknesses in the underlying institutional structures for police

---

[1] *Soto v. State*, 324 N.J. Super. Ct. 66, 351, 1996) (published after the State withdrew its appeal in 1999) concluded that "…unrebutted statistical evidence of disproportionate traffic stops against African-American motorists established de facto policy of targeting blacks for investigation and arrest and thus established selective enforcement violating the equal protection and due process clauses." In Soto, efforts by the State to suppress evidence obtained through race-based stops lead to revelations that New Jersey State Police falsified information to hide pervasive racial profiling and the admission by the State that police engaged in profiling. 734 A.2d 350 (N.J. Super. Ct. Law. Div. 1996); David Kocieniewski, Trenton Charges 2 Troopers with Faking Drivers' Race, N.Y. Times, Apr. 20, 1999, at A23. ("Two state police supervisors said it was common practice for troopers on the turnpike to jot down the license plate number of white motorists who were not stopped and use them on the reports of blacks who were pulled over.")

[2] See Consent Decree, United States v. The State of New Jersey, Civ. No. 99-5970 (MLC), *at* http://www.usdoj.gov/crt/split/documents/jerseysa.htm. The Consent Decree was signed on December 31, 1999.

[3] Id, at provisions 1-5.

[4] See, Jerome Skolnick and Abigail Capolovitz, "Guns, Drugs and Profiling: Ways to Target Guns and Minimize Racial Profiling," 43 *Arizona Law Review*, 413, 419 n. 36 (2001) (reporting that the term "racial profiling" first appeared in the American press was in 1987, followed by a small number of mentions through 1993, then 31 in 1994 and 1995, 63 in 1996 and 1997, 187 in 1998, over 1,000 in 1999, and more than 1,000 again in 2000). A decade earlier, Sheri Lynn Johnson noted the fact of racially selective pedestrian and traffic enforcement. See, Sheri Lynn Johnson, "Race and the Decision to Detain a Suspect," 93 *Yale L.J.* 214, 236 (1983) ("substantial evidence that many police officers believe that minority race indicates a general propensity to commit crime"). By 1999, nearly six Americans in 10 (59%) agreed that "racial profiling is widespread." Nearly four in 10 (39%) agreed that Blacks were "treated less fairly than Whites in dealings with the police." See, Gallup Organization, Americans See Racial Profiling as Widespread, May 13, 2003, available at http://www.gallup.com/poll/8389/americans-see-racial-profiling-widespread.aspx.

supervision and training.[5] Their goal was nothing less than to bring new police technologies into harmony with the constitutional conditions of democratic policing, and to promote police accountability at a time of social conflict between police and minority citizens.[6] Accordingly, the *Soto* consent decree came at a time of great hopes for the reform of policing practices through both constitutional remedies and newly energized political oversight.

The New Jersey consent decree was signed during a period of passion and anger over a series of highly publicized incidents in that state and elsewhere in which African Americans had been subject to egregious incidents of discrimination and police violence.[7] Racial profiling by police on the streets and highways became a prominent feature in the popular and political discourse on policing, rivaling the celebration of the "new policing" that had brought lower crime rates in several cities.[8] This became, then, a moment that promised political accountability and institutional change in law enforcement. Several states initiated data collection procedures

---

[5] Debra Livingston, "Police Reform and the Department of Justice: An Essay on Accountability," 2 *Buffalo Criminal Law Review* 817 (1999); Brandon Garrett, "Standing While Black: Distinguishing Lyons in Racial Profiling Cases," 100 *Colum. L. Rev*. 1815 (2000). See, also, Myriam E. Gilles, "Reinventing Structural Reform Litigation: Deputizing Private Citizens in the Enforcement of Civil Rights," 100 *Colum. L. Rev*. 1384, 1388 (2000), proposing that a law be passed permitting the Justice Department to deputize private suits against police departments for unconstitutional patterns or practices.

[6] David Alan Sklansky, "Is the Exclusionary Rule Obsolete?" 5 *Ohio State Journal of Criminal Law* 567 (2008); Debra Livingston, "Police Discretion and the Quality of Life in Public Places: Courts, Communities, and the New Policing." 97 *Columbia Law Review* 551-672 (1997); Brandon Garrett, Remedying Racial Profiling, 33 *Colum. Hum. Rts. L. Rev*. 41 (2001-2002). But see Samuel R. Gross and Debra Livingston, "Racial Profiling Under Attack," 102 *Colum. L. Rev.* 1413 (2002)

[7] David Kocieniewski, Injured Man Says Brooklyn Officers Tortured Him in Custody, N.Y. Times, Aug. 13, 1997, at B1. "Prosecutors are investigating allegations that police officers beat and tortured a man while he was in custody at a Brooklyn station house."); Associated Press, Violence Erupts Again in Cincinnati, Officers break up protest over police shooting of black man, Los Angeles Times, Apr.11, 2001, at A4. ("The confrontations came four days after Timothy Thomas, 19, was fatally shot as he ran from a police officer trying to arrest him on 14 warrants."); Michael Cooper, Officers in Bronx Fire 41 Shots, And an Unarmed Man is Killed, N.Y. Times, Feb. 5, 1999, at A1. ("An unarmed West African immigrant with no criminal record was killed early yesterday by four New York City police officers who fired 41 shots at him in the doorway of his Bronx apartment building, the police said."); John Kifner, Van Shooting Revives Charges of Racial Profiling by Police, New York Times, May 10, 1998 at B1 (describing incident in which two New Jersey State Police officers fired 11 shots during a traffic stop into a van with four non-white males on their way to a basketball event, wounding three males).

[8] Philip B. Heymann, "The New Policing," 28 *Fordham Urban Law Journal* 407 (2000). These new strategic and institutional designs in policing developed simultaneously with awareness of the prominence of race in policing. See, for example, Jeffrey Goldberg, The Color of Suspicion, N.Y.Times Magazine, June 20, 1999 at 51. See, also, *Washington v. Lambert*, 98 F.3d at 1188 ("'There's a moving violation that many African Americans know as D.W.B.: Driving While Black.' Thirteen Ways of Looking at a Black Man, New Yorker, Oct. 23, 1995 at 59"); Sean Hecker, Race and Pretextual Traffic Stops: an Expanded Role for Civilian Review Board, 28 Colum. Hum. Rts. L. Rev. 551-555 (1997)(discussing the abundance of accounts of racial profiling --- "From the New Jersey Turnpike to the I-95 corridor between Delaware and Florida, empirical studies strongly suggest that police single out minority, particularly African-American, motorists for traffic stops.").

that created an empirical foundation for more extensive analysis of police-citizen interactions, and the Justice Department promoted local data collection efforts, particularly on highway stops and searches, as a way to create dialogue and bring about changes in policy.[9] State and local actors investigated local police practices, with an eye toward possible constitutional violations.[10] The U.S. Justice Department created a module in its National Crime Victimization Survey to develop national population estimates of police-citizen contacts and their outcomes.[11] And, of course, there was litigation in a handful of states, mostly brought by the Civil Rights Division of the U.S. Justice Department. Several were successful and led to Consent Decrees,[12] others resulted in Memoranda of Agreement[13] detailing specific changes in policy and practice, while others led nowhere.[14]

---

[9] Deborah Ramirez, Jack McDevitt, Amy Farrel, A Resource Guide on Racial Profiling Data Collection Systems: Promising Practices and Lessons Learned 84 (2000); Lori Fridell, Racially Biased Policing: Guidance for Analyzing Data From Vehicle Stops (2005), available at http://www.ncjrs.gov/app/publications/abstract.aspx?ID=232491

[10] Elliott Spitzer, "The New York City Police Department's 'Stop and Frisk' Practices," Office of the New York State Attorney General; available at *http://www.oag.state.ny.us/bureaus/civil_rights/pdfs/stp_frsk.pdf (*Results of Investigation of 175,000 Terry stops in New York City, January 1998-April 1999, conducted as part of New York Police Department "Stop and Frisk" Practices)

[11] See, for example, Contacts between Police and the Public, 2005 (available at http://bjs.ojp.usdoj.gov/index.cfm?ty=pbdetail&iid=653). This is the most recent report from the Police-Public Contact Surveys that were conducted every three years from 1996-2005, generating generates epidemiological estimates of the likelihood of a driver being pulled over in a traffic stop and the percentage of all contacts that involve the use of force by police based on a nationally representative sample of more than 60,000 residents age 16. Information includes the reason for and outcome of the contact.

[12] Examples of litigation from 1996 to 2002 that concluded in settlements include: Daniels v. City of New York, No. 99 Civ. 1695 SAS, 2001 WL 62893 (S.D.N.Y.) (Certifying class alleging that the Street Crimes Unit of the New York City Police Department stopped and frisked tens of thousands of predominantly minority New Yorkers without reasonable suspicion); Settlement Agreement, Wilkins v. Maryland State Police, Civ. No. MJG-93-468 (D. Md.) Jan. 5, 1995 (creating a consent decree that required a halt to an alleged practice of racial profiling on I-95 and data collection in all traffic stops) For the text of the consent decrees, see Consent Decree, United States v. City of Los Angeles, Civ. No. 00-11769 GAF *at* http://www.usdoj.gov/crt/split/documents/laconsent.htm; Consent Decree, United States v. City of Steubenville, No. C2 97-966 (S.D. Ohio, Sept. 3, 1997), *at* http://www.usdoj.gov/crt/split/documents/steubensa.htm; Consent Decree, United States v. City of Pittsburgh, No. 97-0354 (W/D/ Pa., Fed. 26, 1997), *at* http://www.usdoj.gov/crt/split/documents/pittssa.htm; In re Cincinnati Policing, Case No. C-1-99-317, U.S. District Court, Southern District of Ohio, Western Division, Collaborative Agreement, April 21, 2002, at http://www.clearinghouse.net/chDocs/public/PN-OH-0005-0008.pdf. For a comprehensive repository of civil rights litigation on policing, see Civil Rights Clearinghouse, University of Michigan School of Law, at http://www.clearinghouse.net/chDocs/public/PN-OH-0005-0008.pdf.

[13] Examples

[14] See, for example, *Chavez v. The Illinois State Police*, 1999 WL 593187 (N.D. Ill. 1999) (dismissing case alleging that thousands of traffic stops throughout Illinois were motivated by race).

In the wake of these and other incidents, Consent Decrees and other instruments designed to bring about institutional reform proliferated despite little empirical evidence of the types of changes that these agreements could most effectively bring about. Because of *Soto's* important place in this history and the broad and deep reach of its remedies, an analysis of the changes that it did and didn't produce opens a unique window to understand the promises and limitations of legal regulation and constitutionalized remedies for racially disparate enforcement by police. Accordingly, we present evidence of race-specific patterns of enforcement in New Jersey a full eight years after the *Soto* Consent Decree was signed, and at a moment when the State claimed success and moved successfully to end federal oversight. We also locate this analysis in the broader scholarly work – legal and empirical - on selective enforcement that has grown over the past decade in lockstep with the growth in court supervision of local law enforcement. Looking ahead, this is a cautionary tale about the design of judicial oversight and monitoring, but also of the resilience of race-based preferences and practices in everyday policing despite the umbrella of oversight and deep institutional changes in the design of policing.

II. Federal Interventions in Local Law Enforcement under Section 14141

A. Federal Litigation as a Strategy for Institutional Reform

The Violent Crime Control and Law Enforcement Act of 1994 was the primary litigation tool in pursuing remedies for law enforcement violations of citizens' rights. The Act, passed in the wake of the 1992 Rodney King Riots in Los Angeles,[15] authorized the Department of Justice to seek injunctions to halt any "pattern or practice of conduct by law enforcement officials" that deprives persons of "rights, privileges, or immunities secured or protected by the Constitution or laws of the United States."[16] The act was not limited to racial discrimination, but could also include police brutality or other types of misconduct. At the time when it was enacted, §14141 was potentially a powerful weapon that permitted the Justice Department to eliminate a wide range of police misconduct—any "pattern or practice of conduct by law enforcement officials" that deprives persons of "rights, privileges, or immunities secured or protected by the Constitution or laws of the United States."[17] Beginning in the mid-1990s, the Department, under §14141, crafted consent decrees that became the leading model for not only for remedying racial profiling, but for the larger project of institutional reform of the police.[18]

As of December 31, 2009, 24 investigations had reached some kind of formal outcome.[19] At least one other – New Orleans – is likely to produce a formal outcome, though at this writing,

[15] 42 U.S.C. § 14141(a) (1994).

[16] Private litigation and litigation to pursue civil rights claims, §1983 actions, and other claims follow separate courses, and are not discussed here. See, for example, Alexander A. Reinart, *Measuring The Success Of Bivens Litigation And Its Consequences For The Individual Liability Model*, 62 Stan. L. Rev. 809

[17] Id.

[18] Garrett, Remedying Racial Profiling, supra note 4.

[19] http://www.justice.gov/crt/split/findsettle.php#CRIPAletters

the timeline is uncertain.[20] No investigation has gone to trial. The complaints under §14141 typically include excessive use of force and racially biased policing.[21]

Outcomes of investigations under § 14141 generally fall into one of three categories: Consent Decrees, Memoranda of Agreement ('MOA'), or Investigative Finding Letters.[22] The content of Consent Decrees and MOAs are similar, as if they had been picked from a menu that was only sporadically updated. They include common reforms and usually, the appointment of a special monitor. There are four common elements, though not all are present in every case: (1) changes in a department's use of force policy, (2) improvements in the citizen complaint process, (3) the creation of an infrastructure for an Early Intervention System (EIS) to monitor officer performance and identify officers whose patterns of interactions with citizens are "inappropriate" or illegal or violate departmental policy, and (4) development of improved training methods and content. Investigative Findings Letters are narrower, generally confined to one of two specific issues, and issuing recommendations without a method to achieve them.[23] As such, they are advisory only, and rarely is an Independent Monitor appointed.[24] Samuel Walker and Morgan MacDonald suggest that the uniformity of the design of these settlements reflects a consensus on "best practices" to promote police accountability that emerged during the first several years of the statute. The consensus is evident in a Department of Justice report, issued in the last days of the Clinton administration in January 2001, that articulated these four prongs of "best practice".[25]

Monitoring was almost always built into MOAs and Consent Decrees, but not Investigative Letters, nor into "stipulated settlements" that result from private litigation.[26] Monitors typically are individuals, firms or *ad hoc* teams of professional consultants. Rarely,

---

[20] CITE LETTER FROM LANDRIEU

[21] The Bush administration favored Investigative Findings Letters over other tools, and a quick audit of the Civil Rights Division website shows a virtual cessation of investigations of law enforcement agencies starting in 2001. See, Walker and MacDonald at 503 and fn 140 for details of the two active investigations launched during the Bush administration.

[22] Multiple outcomes are possible within a jurisdiction. Cincinnati, for example, had both a Consent Decree for some elements of the redesign of policing in that city as well as a MOA. See, In re Cincinnati Policing, C-1-99-317 (2002) U.S. Dist. LEXIS 15928 (S.D. Phio 2002); Memorandum of Agreement between the U.S. Department of Justice and the City of Cincinnati, Ohio and the Cincinnati Police Department (April 12, 2002), http://www.usdoj.gov/crt/split/Cincmoafinal.php. MORE ON LINKAGE

[23] For an example, see Rosenbaum letter to City Attorney in Miami March 13, 2003 (CITE)

[24] See, Samuel Walker and Morgan MacDonald, An Alternative Remedy for Police Misconduct: A Model State Pattern and Practice Statute, 19 George Mason University Civil Rights Law Journal 479 (2008-9).

[25] U.S. Department of Justice, Principles for Promoting Police Integrity: Examples of Promising Police Practices and Policies (2001), http://www.ncjrs.gov/pdffiles/ojp/186189.pdf. See, also, Samuel Walker, The New World of Police Accountability (2005); Brandon Garrett, Remedying Racial Profiling, supra note _.

[26] See, for example, Daniels et al. v. City of New York et al., 99 Civ. 1695 (SDNY) (SAS), http://ccrjustice.org/files/Daniels_StipulationOfSettlement_12_03_0.pdf

courts will monitor directly rather than retain a professional or independent monitor.[27] Monitors generally have prior experience in law enforcement management, in relevant litigation, or in oversight of similar institutional reforms in either public or private domains.[28] Monitors typically review evidence of compliance provided by the state or municipality, and issue reports to the Court and/or to the public on progress toward achieving compliance. They usually work through the termination of the Consent Decree or the MOA. This tenure can be critical when compliance is partial (or worse), since there usually is no obligation attached to either a Consent Decree of a MOA to institutionalize or make permanent any external oversight of the agency. Monitors also are usually restricted to the terms and conditions articulated in the Consent Decree, and cannot exert any jurisdiction or authority if they become aware of other problems such as employment discrimination or corruption.[29]

The duration of monitoring in the first wave of Consent Decrees and MOAs before 2001 was usually set at five years, though several were extended based on evidence of either noncompliance or limited compliance.[30] New Jersey's, as we discuss below, was scheduled to expire after seven years, but was extended to 2009 while transitional activities took place.[31] The monitoring function can be terminated early if a Court terminates a Consent Decree before its scheduled expiration.[32]

The design of monitoring in many of the first wave of Consent Decrees and MOAs placed heavy responsibility on the monitors, and were quite optimistic about their capacity to work as change agents to bring about institutional reform. Their roles can vary and overlap, including collaborators, technical advisors, auditors, "police" to the police, shaming agents, analysts, or stern parents. Their reports, often containing detailed descriptions of the internal design and workings of a law enforcement agency, have the ancillary benefit of shining light on what often are insular if not closed agencies that resisted any outside inquiry about policy and procedure.[33] They also served to advance the meta-goal of enhancing accountability that was at the heart of both the passage of § 14141 and of the early consent decrees.[34] By opening police agencies to new forms of democratic oversight, the intervention of federal monitors had the aspiration that consent decrees would transcend their formal procedural aims to penetrate the

[27] Example

[28] See, Section XX, infra, and Appendix A.

[29] Walker and MacDonald, supra note _ at 511. Example.

[30] LA, Ohio

[31] Infra Section II.B.

[32] Example in Cincinnati

[33] See, for example, Jerome Slolnick and James J. Fyfe, Above the Law (1994); CITES

[34] Debra Livingston, Debra Livingston, Police Reform and the Department of Justice, supra note 5. QUOTE

informal organizational cultures that shapes officers' everyday conduct on the streets and highways.[35]

B. Racial Profiling in New Jersey and the 1999 Consent Decree

In December 1999, the U.S. Department of Justice and the State of New Jersey entered into a consent decree to oversee and monitor the enforcement activity of the New Jersey State Police in its actions on the state's highways[36]. The consent decree followed the holding in *Soto v. State* that the New Jersey State Police had engaged in a pattern and practice of racially selective enforcement on the New Jersey Turnpike.[37] Other claims of constitutional violations, including excessive use of force, took a back seat to the claims of racially selective enforcement by NJSP troopers in the selection of vehicles to be stopped and in the subsequent selection of vehicles to be searched.[38]

The *Soto* consent decree was unique among the early consent decrees enacted in the U.S. in the 1990s. Other consent decrees signed in that decade involved claims of civil rights violations against municipalities whose police departments were accused of systematic patterns of excessive force and other wrongdoing.[39] New Jersey was the first to name a state agency, the New Jersey State Police. It put into place a detailed plan for institutional reform in the recruitment, training, and supervision of NJSP troopers. The Consent Decree also required the design and implementation of management information system to compile data on the patterns of enforcement and the outcomes of vehicle stops and searches.

---

[35] See, for example, Barbara Armacost, Organizational Culture and Police Misconduct, 72 *George Washington Law Review* 453, 509 (2004) (noting the distinction between procedures aimed at "rotten apples" and those aimed at "rotten barrels").

[36] Consent Decree, United States v. The State of New Jersey, Civ. No. 99-5970(MLC). Full text available at: http://www.nj.gov/lps/jointapp.htm

[37] *Soto v. State*, 324 N.J. Super. 66, 351, 1996) (approved for publication after withdrawal and dismissal of the State's appeal on April 22, 1999) (holding that "…unrebutted statistical evidence of disproportionate traffic stops against African-American motorists established de facto policy of targeting blacks for investigation and arrest and thus established selective enforcement violating the equal protection and due process clauses.")

[38] The *Soto* court had found that the State and the NJSP troopers were engaged in racial profiling to the extent that they were stopping African American motorists on the New Jersey Turnpike at higher rates than whites for speeding violations. *Soto, id.* Peter Verniero, Attorney General, Interim Report of the State Police Review Team Regarding Allegations of Racial Profiling (1999) (stating that four of ten stops were of minorities and eight of ten searches, "the overwhelming majority," were of minorities). See also David A. Harris, The Stories, the Statistics, and the Law: Why "Driving While Black" Matters, 84 *Minn. L. Rev*. 265, 277-289 (1999) (reviewing the first comprehensive statistical analysis of police stops and race in New Jersey, Maryland and Ohio).

[39] Brandon Garrett, Remedying Racial Profiling, supra note _. Debra Livingston, Police Reform, supra note _.

The Consent Decree required that the NJSP implement new policies and extensive reforms in the training and supervision of NJSP troopers, and implemented careful external monitoring both of compliance with procedural and substantive reforms. Compliance also was determined by progress toward the elimination of racial disparities in the patterns of NJSP vehicle stops and searches on the New Jersey Turnpike that animated the *Soto* litigation.[40] The court-appointed Monitor overseeing the Consent Decree for the U.S. District Court published a series of 16 reports between ___ and August 2007 documenting the State's compliance with the stipulations of the agreement. Report No. 16, the penultimate report issued in August 2007, stated that the State and NJSP had made substantial progress in reducing racial disparities in the conduct of vehicle stops, and that the State and the NJSP were in full compliance with the terms, conditions, policies, procedures, standards and benchmarks of the Consent Decree.[41]

The State was so confident that its compliance record would persuade the Court to end its supervision that it began planning for the post-Consent Decree institutional design of the reformed State Police even before the Monitor's 16th report in August 2007.[42] In public hearings in September 2007[43], in the Monitor's August 2007 report,[44] the State claimed that the time had come to lift the burden of federal oversight from the New Jersey State Police. A bipartisan Governor's Advisory Committee's on Police Standards and Practices agreed, and echoed the conclusion of the Monitor in its final report in December 2007.[45] The state's law enforcement institutions reported to the Advisory Commission, quoting the Monitor, that the State had the

---

[40] See, __ for details on the specific conditions and requirements set forth in the consent decree, at http://www.nj.gov/lps/jointapp.htm

[41] Monitors' Sixteenth Independent Report, Long-Term Compliance Audit, August 2007, U.S. District Court, District of New Jersey, available at: http://www.state.nj.us/lps/monitors-report-16.pdf (concluding that: "Compliance requirements are now at 100 percent levels. Policy, training, supervision, inspections and audits….are fully staffed, fully functioning, and, in the opinion of the monitors, fully capable of self-monitoring and self-adaptation.").

[42] An expert report commissioned by the New Jersey State Attorney General and submitted eleven months before the Monitor's August 2007 report, foresaw the Monitors' conclusions on compliance, and issued recommendations that post-Consent Decree monitoring of the NJSP practices be assigned to the State Attorney General. See, Samuel Walker, Memorandum to the Attorney General, State of New Jersey, "Post-Consent Decree Oversight of the New Jersey State Police" September 16, 2006 (concluding that "[t]he NJSP is to be commended for successfully implementing the terms of the Consent Decree in a timely fashion.") (available from authors).

[43] http://www.state.nj.us/acps/home/hearings/070924.html

[44] *Supra* note __, at iv. In its 16th and penultimate report, the Monitor noted that "the State Police had become 'self-monitoring' and that as agency it possessed the ability to "analyze and correct" problematic law enforcement procedures on a 'real time' basis." This was seen by the Advisory Committee in its December 2007 Final Report as reliable evidence that the state had created an infrastructure that was capable of interalizing the oversight responsibilities that had, under the Consent Decree, been assigned to the court-appointed Monitor.

[45] New Jersey Advisory Committee on Police Standards, Report and Recommendations to Governor Jon S. Corzine Pursuant to Executive Order No. 29, December 7, 2007, available at: http://www.state.nj.us/acps/njacps_final_report.pdf (recommending that "the State join in the motion to dismiss the Consent Decree").

political will and the infrastructure to assume responsibility for ongoing oversight and monitoring of the practices of the NJSP.[46] In the 17th and final report before the state was released from the consent decree, dated April 2009, the Monitor reiterated the claim from its penultimate report that "the State Police had become 'self-monitoring' and that as agency it possessed the ability to 'analyze and correct' problematic law enforcement procedures on a 'real time' basis."[47] The legislature passed the ___ Act in ___, transferring monitoring and oversight of the NJSP to the Office of State Police Affairs within the Office of the State Attorney General. The parties to the Consent Decree, the State of New Jersey and the Civil Rights Division of the U.S. Department of Justice, filed a joint motion in August 2009 to dissolve the order.[48] On September 20, 2009, Judge Mary L. Cooper of the U.S. District Court in Newark signed an order dissolving the Consent Decree.[49]

C. The End of Profiling?

The independent Monitor, as well as the Attorney General's expert, focused their reports on policy, procedure and practice, rather than addressing the central issue in *Soto*: whether NJSP troopers engaged in race-based selection of vehicles and drivers for stops and potentially searches.[50] When the Monitor last visited the question of race-specific patterns of enforcement in

---

[46] Pinpoint on this in Adv Comm Final Report.

[47] Cite 17th Monitor Report, April 2009.

[48] CITE

[49] Oversight of New Jersey State Police is Ended, *New York Times*, September 21, 2009, at http://www.nytimes.com/2009/09/22/nyregion/22profile.html.

[50] Neither the Monitor's reports nor the AG's Expert Memorandum addressed the basic "supply-side" question that was litigated in *Soto* and that animated the Consent Decree: the racially disparate practice of selection of motorists for stops by NJSP troopers on the southern portion of the Turnpike. In fact, none of the monitoring reports addressed the claims of racial disproportionality in the selection of vehicles for stops, the primary concern that animated *Soto* case and the Consent Decree.

Yet others did, and offered credible evidence of racial disparities in the selection of vehicles for stops, relative to estimates of the base rate of moving violations. See, John C. Lamberth and Joseph B. Kadane, *In the Matter of the Study of State Police Stop Activity at the Southern end of the New Jersey Turnpike*, http://www.state.nj.us/acps/home/hearings/pdf/061121_kadane-lambert.pdf (hereafter, LK Report). The LK Report focused on vehicle stops on the section of the New Jersey Turnpike patrolled by the Moorestown Station, between Exits 1 and 7A, the area that was the focus of the *Soto* opinion. The LK report analyzed motor vehicle stops by New Jersey State Police ("NJSP") troopers on the southern end of the New Jersey Turnpike in August and September 2005. LK estimated the distribution of the racial and ethnic identity of drivers who exceeded the speed limit, and compared this distribution to the distribution of drivers who were stopped by New Jersey State Police troopers in the same sections of the New Jersey Turnpike. The LK Study concluded that NJSP troopers were more likely to take enforcement action against black violators than they were against white violators. Using two different measurement methods, they estimated that between 18.5 and 19.0 percent of African-American motorists committed speeding violations between Exits 1 and 7A of the Turnpike and were eligible for stops. Their analysis of NJSP data on vehicle stops for speeding in the same area showed that 30.5 percent of those stopped for speeding were African American. Tests showed that these

Report No. 16,[51] he estimated the extent of racial disparities in the outcomes of stops: searches, deployment of canine units, and arrests. The analysis was based on a sample of 269 "critical event" stops made by NJSP Troops B and E in the second half of 2006.[52] No information was given on the definition of a "critical event" stop, the method for selecting these cases, or how these cases stacked up against the full universe of over 12,200 stops during that time period.[53]

Nevertheless, Report No. 16 was cited by the State and the Governor's Advisory Committee as evidence of compliance in its recommendation to seek termination of the Consent Decree. The report concluded that there was no evidence of racial disparity in the conduct of NJSP troopers pursuant to stops, despite showing statistically significant differences in the rates at which drivers were asked to consent to searches[54] and the rates at which canine units were deployed for vehicle searches.[55] The Monitor minimized the importance of racial disparities in post-stop decisions by citing "qualitative" factors that may explain the differences.[56] Qualitative factors were defined as "levels of discretion" which in turn were equated with reasons for the stop or the nature of the violation. These were categorized as high, median (sic) or low discretion events. The Monitor used these categories to classify cases, and then re-analyzed the

---

differences were statistically significant. See, also, John Lambert and Joseph B. Kadane, Are Blacks Egregious Speeding Violators at Extraordinary Rates in New Jersey? 8 *Law, Probability and Risk* 69 (2009).

The LK Report was independently reviewed in 2007 at the request of the Governor's Advisory Commission. Jeffrey Fagan et al., Memorandum to James E. Johnson, *Lamberth-Kadane Report on New Jersey State Police Stop Activity on the Southern End of the New Jersey Turnpike*, April 10, 2007, available at: http://www.state.nj.us/acps/home/hearings/pdf/dpny_22440631_v1_peerreviewteam.pdf. The Panel also examined data on stops and searches during 2005, and included a second window in 2005 (March-April) as a validity check against seasonality. The Panel reported that there was no social science basis to reject the LK Study, and that the conclusion that racial disparities were present in the selection of drivers for stops by NJSP troopers was valid. While the Panel questioned the level of the disparity between black and white stop rates reported in the LK Study, it affirmed the likelihood that a racial disparity existed. The Panel concluded that after taking into account any limitations in design and measurement introduced uncertainty in the estimates computed by LK, and the confidence intervals around the estimates of moving violation rates by Blacks would still suggest that stop rates for Blacks are disproportionate to their violation rates and disproportionate to the rates for drivers of other races.

[51] Monitors' Sixteenth Independent Report, supra note__ ,at 8.

[52] Id. Troop B patrols state highways in a broad area of northern New Jersey, which encompasses primarily rural counties. rural areas to the north and west of those population centers. Troop E patrols includes three substations whose patrol areas include portions of the New Jersey Turnpike. The Moorestown Station, which was the focus of the Soto litigation, was not included in the Monitors' analysis.

[53] Based on data made available to the authors by the Office of the Attorney General of New Jersey. We used G*Power to estimate the power of a sample of 269 cases and an effect size of .25 for two of the three groups. Power is less than .85.

[54] Monitors' Sixteenth Independent Report, supra note __ at 14-16.

[55] *Id*., at 16.

[56] *Id*., at 20-21

data on post-stop outcomes to determine the extent of racial disparities conditional on these factors.

Low discretion stops were defined as those "activities that *will almost always* result in a law enforcement response (emphasis in original)". Examples of "low discretion" events include suspicion of criminal activity, driving under the influence, reckless driving. Median discretion events are those that "*usually*" result in a law enforcement response, such as following too closely, aggressive driving, or undocumented vehicles (e.g., expired registration or license or insurance documentation). High discretion stops include equipment violations such as cracked windshields or broken taillights, or other non-speeding moving violations such as improper lane changes. Using this nosology, the Monitor concluded that there was, in fact, no evidence of racial discrimination.

But the evidence for this conclusion leaves open several questions about its reliability and accuracy. First, the *post hoc* classifications were based on and limited to information available to the Monitor that was unobservable in the classification exercise. There is no reasonable way to conclude from the details of the stop reported in the administrative (observational) database whether the "guilty characteristic" that motivated the initial action was present and in what form. The classification is based mainly on offense categories, not on any behavioral indicia that might impart "suspicion" to the trooper. The robustness of the conclusion depends on the sensitivity of the classification scheme. Even small disagreements and reclassifications could shift the results, given the small sample sizes. Given the importance of "suspicion" as a component of the "high discretion" category, and the racial dimensions of suspicion-based actions by legal actors, race cannot be excluded as a factor in decisions to search in those discretionary circumstances.[57] Post-hoc outcomes tests of taste-based discrimination are unsatisfactory precisely because of the unknown distributions of the unobservable indicia of suspicion that signal the guilt of the individual and that might satisfy a preference to search one set of persons over another.

Second, the sample of stops that did not proceed to a second order interaction were not tested, so censoring is potentially a serious problem. Searches were requested in 146 of the 269 cases reviewed, and there is no information provided on the rest. Yet, since the distribution of both the observable and unobservable characteristics in the stopped population is censored in the analysis, any estimates of the racial distribution of post-stop activities has unknown error rates and likely are biased.[58] In fact, there is no information available in the administrative dataset beyond the rationale for the stop. Accordingly, the monitors' report offers no comparisons that directly estimate the likelihood of a search within the total sample of vehicles and drivers and that allow for control of the observables.

---

[57] See, R. Richard Banks et al., Discrimination and Implicit Bias in a Racially Unequal Society, 94 *Cal. L. Rev*. 1169 2006 ___, for a summary. See, also, R. Richard Banks, Beyond Profiling: Race, Policing and the DrugWar, 56 *Stan. L. & Policy Review* 571 (2003). See, also, Geoffrey Alpert et al., XXX; Sandra Graham and Brian Lowery, XXX;

[58] A quick look at the data in the Monitors' Report No. 16 shows that Black and White drivers were searched in about one in three "high discretion" cases in three where a search is requested, and Hispanic drivers are searched in nearly half the cases (47.3%). Canine units are more than 13 times more likely to be deployed in "high discretion stops" of Black and Hispanic drivers compared to White drivers. Monitors' Sixteenth Independent Report, supra note __ at Tables 6 and 7.

Third, the classification method assumes a constant motivation of discrimination, and interprets the results accordingly. But not all encounters are motivated by a taste for discrimination; not only does the information available to police vary by incident, but so too do the motivations for engaging citizens initially, and then to continue the engagement through additional activities. In fact, there is some evidence that when the police officer's motivation is to minimize crime, there is no evidence of racial preferences.[59] The motivations could be race-neutral yet still produce racial disparities, assuming that there is some racial disequilibrium either in the indicia of suspicion or in the base rates of illegal behavior.[60] This could, in fact, be simply statistical discrimination, however unethical it may be to simply "play the base rates,"[61] or it could be race-based preferences, or a race-based skew how "suspicion" is interpreted. This heterogeneity in preferences interacts with unobservables to make the story more complicated than is assumed in a single motivation model. In effect, unobservables shape officers' preferences to search both actual offenders and potential offenders (regardless of guilt).[62]

D. Discrimination During and After the Consent Decree

How then are we to understand the Monitor's evidence and interpret the claims that NJSP officers no longer discriminate? In his 16th report, the Monitor concludes that there are significantly higher rates of searches and canine deployments for African Americans who were stopped, but minimizes their importance by citing "qualitative" factors that may explain the differences. The Monitor avoids a conventional test that uses a common metric for identifying disparate treatment of persons in a protected group – comparison of their outcomes with the outcomes of persons who are "similarly situated" but in a different group – to assess whether non-Whites selected for stops by police are subject to equal probabilities of selection for further interactions.[63]

The general test for evidence of disparate treatment is a regression equation that takes the form:

$$\text{Outcome} = \alpha + \beta_1 * \text{Minority} + \Sigma_i \beta_i * (\text{Plausible Non-Race Influences}) + \varepsilon,$$

[59] Dominitz and Knowles

[60] Knowles, Persico and Todd, 2001; Persico, 2009

[61] See, for example, Bernard Harcourt, Against Prediction (2006); But see, Risse and Zeckhauser, (2004), who defend the practice.

[62] Brock et al., On the Observational Implications of Taste-Based Discriminaliton in Racial Profiling (mimeo). See, also, S. Anwar and H. Fang, "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence, 96 American Economic Review 127-151 (2006); K. Antonovics, and B. Knight "A New Look at Racial Profiling: Evidence from the Boston Police Department, 91 Review of Economics and Statistics163-177 (2009).

[63] See, an illustration, Ian Ayres and Jonathan Borowsky, A Study of Racially Disparate Outcomes in the Los Angeles Police Department, at 5 (October 2008), available at http://www.aclu-sc.org/documents/view/47."

where *Outcome* is the event or status of interest, *Minority* is an indicator for the racial composition or status of the unit observed (i.e., precinct or person, depending on the outcome), *Plausible Non-Race Influences* are a set of variables representing non-race factors that also might influence the outcome, and an error term ε that captures the variation in the outcome that cannot be explained by either Minority status or the Non-Race Influences. These models may include non-race influences that are correlated with race, so as to better identify the unique effects of race that are present once the influence of proxies for race are removed.[64]

Consider the following example, from *Griggs v. Duke Power Co.*, an employment discrimination case.[65] In a disparate treatment claim, one could test whether the use of a high school diploma requirement biases the hiring process since African American job applicants may be less likely to have obtained a high school diploma. Had this race-correlated control is introduced, it would likely have reduced the racial disparity in the hiring rates – for the simple reason that minority applicants at that time were less likely to have obtained a high school diploma. Should a statistical test control for whether or not an applicant had a high school diploma? As Ian Ayres points out,[66] in a disparate treatment case, the answer is yes. Under a disparate treatment theory, the critical question is whether an applicant's race was the cause of being denied employment. If applicants were rejected because the employer chose not to hire diploma-less applicants, the applicants' race would not be a "motivating factor" in employer's decision. The goal in specifying these models is to identify the effects of race on outcomes after simultaneously considering factors that may be relevant to race.[67]

We conduct precisely such an analysis in this paper. We conduct a series of tests on the likelihood of a search or a seizure of contraband pursuant to a search. We include observables based on characteristics of the vehicle and the driver, as well as the nature of the violation. We make no assumptions about preferences or tastes for discrimination, only that there is heterogeneity in that preference among officers. Following Close and Mason[68] and Anwar and

---

[64] For a general discussion of the specification of regression models to test for disparate treatment, see generally D. James Greiner, "Causal Inference in Civil Rights Litigation," 122 *Harvard L. Rev.* 533 (2008). For a general discussion of how regressions sort out the influences of predictors of an outcome, see Thomas J. Campbell, Regression Analysis in Title VII Cases: Minimum Standards, Comparable Worth, and Other Issues Where Law and Statistics Meet, 36 *Stanford L. Rev*. 1299 (1984).

[65] *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971). Ian Ayres uses this case to illustrate the difference between a disparate impact claim, which has limited value in an employment discrimination claim, and disparate treatment. See, See, e.g., Ian Ayres, "Testing for Discrimination and the Problem of 'Included Variable Bias'," Yale Law School Working Paper (2010), available at http://islandia.law.yale.edu/ayers/ayresincludedvariablebias.pdf

[66] Ian Ayres and Jonathan Borowsky, A Study of Racially Disparate Outcomes in the Los Angeles Police Department, at 5 (October 2008), available at http://www.aclu-sc.org/documents/view/47."

[67] See, e.g., Ian Ayres, "Testing for Discrimination and the Problem of 'Included Variable Bias'," supra note _. See, also, Ian Ayres, "Three Tests for Measuring Unjustified Disparate Impacts in Organ Transplantation: The Problem of "Included Variable" Bias," 48 *Perspectives in Biology and Medicine* 68 (2005)

[68] Cite

Fang,[69] we assume that crime minimization is the motivation for policing, but that unobservables make it difficult to attribute differentials in searches to a preference for discrimination. Similar to others estimating taste-based discrimination,[70] we exploit the observed heterogeneity in police officer race to establish search and success rates of suspects of different races across officers of different races. Since the tests are sensitive to assumptions about the purpose of the search – crime minimization versus punishment maximization, for example – we focus our analyses on the decision to search conditional on a stop.[71]

This also makes sense in light of the institutional contexts in New Jersey, where there was "smoking gun" evidence of a preference to discriminate.[72] To estimate the influences of local NJSP contexts and cultures, given these institutional preferences, we extend our estimates to take into account the unique effects of each organizational unit of the NJSP.

III. Theory and Research on Race and Selective Enforcement

The term "racial profiling" is a broad descriptive category that encompasses a range of linked practices by police and other legal actors. It may include the use of race as a single or dominant factor in the decision to stop a person or car, and once stopped, to frisk a citizen, to conduct a search of her person or vehicle or passengers in the vehicle, to use noxious and intrusive methods for search such as canine deployments, or to use physical force or weapons to prevent escalation or aggressive reactions by citizens who have been stopped or detained. So, understanding the role of race in these encounters requires that we consider the role of race not just in the initial encounter, but in judgments on whether and how to proceed beyond the initial encounter.

Police engage in racial profiling when they select persons of a specific race for attention because they assume that those persons are more likely to commit a targeted crime or crime generally than a white or majority person.[73] The decision to stop a person, then, reflects an *a*

---

[69] Anwar and Fang, supra note _.

[70] Anwar and Fang, Close and Mason, David Bjerk

[71] Dominitz and Knowles, for example, note that assumptions of equilibriums or equal search success rates provide a better fit when the goals of enforcement are punishment maximization instead of crime minimization. But given unobservables, the decision about how to model becomes a probabilistic notion with unknown contingencies. At the same time, a finding of equal search rates doesn't necessarily rule out either motivation, nor does it rule out the existence of racial prejudice.

[72] See, Soto v State, supra note 1. See Veneiro memo, supra note 13.

[73] See, for example, International Association of Chiefs of Police [IACP], 1999; U.S. General Accounting Office, 2000). Skolnick and Caplovitz reduce profiling to its core racial component: the selection of individuals for heightened scrutiny based on the color of their skin. See, Jerome Skolnick ___. Jeffrey Fagan et al. show the probability of selection by police for heightened attention (e.g., stop, frisk, search) is conditional on both race and neighborhood. *See*, Jeffrey Fagan et al., *Street Stops and Broken Windows Revisited: Race and Order Maintenance Policing in a Safe and Changing City*, in

priori probabilistic assessment by police officers that members of a specific race are more likely than "similarly situated" members of another race to be engaged in criminal activity, and therefore select them for stops.[74] But the decision to proceed further may also reflect probabilistic judgments about race, or about other factors conditional on suspect race.

Evidence of racial disparity in post-stop outcomes suggests that these probabilistic assessments are hardly limited only to the first stage of an encounter, when an individual is initially selected for attention. Racial profiling can occur "when a police officer stops, arrests, questions, searches or otherwise investigates a person because the officer believes that members of that person's racial or ethnic group are more likely than the population at large to commit the sort of crime that the officer is investigating".[75] So, there may be race-based probabilistic assessments inherent in the decision to frisk a person stopped on the street or to search a vehicle once that vehicle has been selected. Perhaps the indicia of suspicion and their algebra change from one stage to the next: some indicia that are present in selection for the initial encounter are no longer relevant in the decision to proceed further, while other indicia that were not factors in the initial decision to launch an encounter enter into the decision logic in later stages. For example, speeding may strongly influence the decision to stop a vehicle, but the speeding may have less (if not nothing) to do with the decision to search a vehicle. [76] Yet race may influence both the decision to stop the vehicle and to proceed further to a search. Demeanor also may signal suspicion, and the subjective interpretation of demeanor as "suspicious" or masking

---

EXPLORING RACE, ETHNICITY AND POLICING: ESSENTIAL READINGS (S. Rice and M. White, eds.) (2010).

[74] Whether racial profiling includes stops resulting from race-based suspect descriptions is controversial. Gross and Barnes (2001), Rudovsky (2001), Cole (1999), Kennedy (1997), Davis (1997), Durlauf (2006) and several others carefully distinguish such stops from stops based on discretionary judgments about suspicion. Case law also distinguishes racial profiling from police stops motivated by race-based suspect descriptions. In *Brown v. Oneonta* (221 F.3d 769, 2nd Cir. 1999), for example, the Court of Appeals in the 2nd Circuit found no evidence of an equal protection violation when police stopped over 200 African- American males (and one woman) in the City of Oneonta based on a suspect description provided by a victim of attempted rape. But Richard Banks claims that race-based suspect descriptions are no different than other stops motivated by racial classifications. Richard R. Banks, *Race-Based Suspect Selection and Colorblind Equal Protection Doctrine And Disclosure*, 48 UCLA L. REV. 1075 (2001) (showing that by virtue of the similarity of race-based suspect descriptions and race-based profiles, both offend the principles of colorblindness in the Equal Protection Clause in that they disproportionately burden racial minorities, and both should be subject to strict scrutiny within the Fourteenth Amendment). Banks cites four weaknesses in victim-supplied suspect descriptions that reduce the targeting of suspects to the same probabilistic (and therefore flawed) determination as in a broader race-based profile: (1) the fallacy of appearances of a category of people who resemble the suspect, (2) the unreliability of victim descriptions, (3) the overly broad application of the suspect description, and (4) the excessive weighting of race in a "multiple factors" suspect description.

[75] Samuel R. Gross and Deborah Livingston, Racial Profiling Under Attack, _ Colum L Rev __ (2002). See, also, Samuel R. Gross and Katherine Barnes, Road Work__, __ Mich. L. Rev. __ (2001)

[76] Similarly, in a pedestrian stop, the decision to stop a suspect may reflect subjective judgments about behaviors or other indicia that are correlated with crime, but the decision to search that suspect may be driven by other factors. However, the constitutional regulation of the street stop versus the vehicle stop are quite different. MORE

illegality or signaling culpability may itself be subject to racialized interpretations.[77]

The linkage between these two probabilistic assessments shows that race and suspicion are linked and interact across the stages of an event. These interactions are events, and each decision in an event is conditional on the preceding one yet linked by race and perhaps other factors. While that linkage may vary from one stage to the next, the evidence of racial disparities both in the selection of cases and what happens following selection suggests the persistent infusion of race in the sequence of police-citizen interactions.

Yet researchers studying racial profiling or racially selective enforcement have consistently disaggregated this selectivity into two stages of the police-citizen interaction: the initial selection of individuals, and the ensuing stages of interaction where police may decide to end an encounter or go on to further and more intrusive interdictions. Compare, for example, the analysis by Andrew Gelman and colleagues on the racial distribution of pedestrian stops in New York City[78] with the analysis by John Knowles, Nicola Persico and Petra Todd of searches and "hit rates" pursuant to vehicle stops Maryland.[79] Gelman and colleagues analyze pedestrian stop rates in New York City by neighborhood, taking into account the crime conditions and racial composition of the area, and find racial disparities in stop rates. Although they acknowledge that these disparities may persist in post-stop outcomes, they don't engage the question of how race links from one stage to the next. Knowles and colleagues do not engage the question of the production of stops eligible for search.

To some extent, the separation is understandable. Different legal, policy and sociological questions animate these two sets of studies, for better or worse. For example, …..Data availability also presents challenges to establishing a linkage between base rates of "suspicious" behavior and subsequent contacts with police, as does reliable measurements to estimate race-specific base rates of behaviors eligible for police interdiction.[80]

MORE

IV. Data and Methods

We test for racial bias in motor vehicle searches conducted by the New Jersey State Police over an 18-month period from 2005-7, the final years of the 1999 *Soto* consent decree. We adopt an analytic framework based on recent work by Close and Mason that exploits

[77] Cites on race and suspiciousness…. Graham, Bridges, Goff, Eberhardt, Banks on profiling research

[78] Gelman, Fagan and Kiss

[79] J. Knowles, Nicola Persico, & Petra Todd, Racial bias in motor vehicle searches: Theory *and evidence*. JOURNAL OF POLITICAL ECONOMY*, 109,* 203–299 (2001)

[80] Even reliable measurement may fall short in explaining stops and their outcomes, given the incidence of either pretextual or suspicionless stops. See, >>>

heterogeneity in officer and suspect race.[81] We test the null hypothesis that neither an officer's nor a suspect's race predicts the probability of a search or seizure, and in turn that no combination of officer and suspect race predicts a search.

A. Data

The NJSP provided incident-level records on 257,059 stops, searches and seizures on the New Jersey Turnpike from October 31, 2005 through March 31, 2007. The data were issued on request of the *Advisory Commission*, as part of its ongoing assessment of compliance by the NJSP with the terms of the consent decree. The data are part of the MAPPS[82] system developed as one of the obligations of the consent decree. MAPPS records data generated by troopers pursuant to each stop. Elements include demographics of the driver (but not passengers), information about the vehicle stopped, the reason for the stop (moving or non-moving violation), and all post-stop interactions (searches, arrests, use of force). The data are collected for each trooper, including the trooper's demographics, years of service, rank, and unit of deployment.

Dependent variables are search (binary) and seizure (binary). Search included those conducted based on consent request, searches incidental to arrest, and probable cause searches. Searches were conducted in 10,890 stops (4.24%), under several rationales, both constitutional and extra-constitutional. Although searches incidental to arrest were banned in New Jersey following *State v Eckel*, [83]13.2% of the searches in the data were conducted under this rationale, generally in stops predating the decision. Most (47.6%) were recorded as probable cause searches. Consent searches were 3.7% of all searches, despite the ban on consent searches following *State v. Carty*[84]. However, there is considerable missing information in the rationale for the search, which precluded tests for racial differences by search rationale.

Seizure was extremely rare: 0.52% of all stops. Seizure was defined as a stop where there was any seizure of contraband. Details of the seizure in terms of type (e.g., weapons or drugs) or amount/value were not made available. Arrests were made in 3.8% of all stops. We did not analyze arrest as an outcome given its ambiguous meaning relative to the stop itself. Some arrests were made pursuant to behavior after the stop, such as resisting arrest or assault on a trooper, so we were unable to determine which arrests were made as a result of the detection of contraband or other law violations as a result of the stop.

Predictors included race of driver and trooper, state license plate, trooper years of service, and the stop rationale – moving violation or other violation. For moving violations, information

---

[81] Cite

[82] Management Awareness of Personnel Performance. See, Consent Decree, supra note __, at § 40-56

[83] 374 N.J. Super. 91, 2006

[84] 790 A.2d 903, 905, 907 (N.J. 2002). See, also, George C. Thomas III, The Short, Unhappy Life of Consent Searches In New Jersey, 36 Rutgers Law Record 1 (2009)

on the precise speed was unavailable.[85] We assume that out-of-state license plate would be a marker of suspicion on an interstate highway, and especially on the Interstate 95 corridor through the urban megalopolis of the Northeast, as a potential drug courier and prompt a search.

B. Identification Strategy

Following Close and Mason,[86] we test the separate and interactive effects of trooper and driver race on the treatment of drivers in the period nearing the end of the *Soto* consent decree. While Close and Mason examined driver-officer racial pairings non-parametrically, we test for interactions using a regression framework.

We first test whether the probability that a traffic stop leads to a search (S_i) varies by driver race. For each driver *i*, we estimate Model 1:

$$Logit(SEARCH_i) = \alpha + \sum_{race\ group\ r=1}^{R} \beta_r \boldsymbol{DriverRace}_i + \varepsilon$$

where $DriverRace_i$ is a vector of dummy variables indicating the race of driver *i*. We expand this model to Model 2, which includes covariates $\boldsymbol{X}$ describing the circumstances of the stop (whether the driver is from New Jersey, New York, Pennsylvania, or elsewhere, whether they were stopped for a moving or nonmoving violation, etc.)

$$Logit(SEARCH_i) = \alpha + \sum_{driver\ race\ r=1}^{Total race\ groups\ R} \beta_r \boldsymbol{DriverRace}_i + \boldsymbol{\gamma X}_i + \varepsilon$$

---

[85] The precise speed might be important to determine the probability of a driver being stopped, but its relevance to the probability of being searched is indirect at best. Assume that a driver whose speed exceeds the legal limit by 15 miles per hour is an egregious speeder. This presents a low discretion situation, and the person is more likely to be stopped than would a person traveling at less than five miles per hour above the speed limit. That speed is a high discretion situation. Persons traveling at speeds in between, or 5-15 miles per hour over the speed limit, can reasonably be classified as presenting moderate discretion to the detecting trooper. See, e.g., Joseph B. Kadane and John Lamberth, Are Blacks Egregious Speeding Violators at Extraordinary Rates in New Jersey? 8 *Law, Probability & Risk*, 139-152, 2009. In their analysis of stops on the New Jersey Turnpike, they reported that the percentage of Black egregious speeders was only slightly greater (19.0%) than the percentage of Black speeders overall. There is no reason a priori to assume that excessive speed signals suspicion to a trooper and thus increase the likelihood of a search. Excessive speeding would draw unwanted law enforcement attention, perhaps the last thing that a drug courier on an interstate highway (especially I-95) would welcome. If we assume that drug couriers are sufficiently rational to adjust their illegal behavior in light of risk and detection probabilities to reach an equilibrium with other drivers, then they are also rational (and smart) enough to avoid attracting attention by excessively speeding.

[86] Close and Mason, supra note __.

In Model 3, we test whether search probability varies not only by driver race, but also by the race and the years of service of the officer making the stop. In each of these models, we include fixed effects and a linear time trend for each calendar quarter in the period. We also include fixed effects for the officer's organizational assignment to account for the distinct regions of the state that officers patrol, and to account for the culture and history of each station.[87] Because of the importance of the Moorestown station in the *Soto* litigation and evidence,[88] we then estimate Models 1-3 separately for stops made by troopers assigned to that station.

$$Logit(S_i) = \alpha + \sum_{driver\ race\ r=1}^{Total\ race\ groups\ R} \beta_r \boldsymbol{DriverRace}_i + \sum_{trooper\ race\ t=1}^{Total\ race\ groups\ R} \beta_t \boldsymbol{TrooperRace}_i + \gamma X_i + \varepsilon$$

To test the extent to which driver and trooper race interact in the determination of stop outcomes, we then estimate Model 4:

$$Logit(S_i) = \alpha + \sum_{driver\ race\ r=1}^{Total\ race\ groups\ K} \sum_{trooper\ race\ t=1}^{Total\ race\ groups\ K} \beta_{rt} \boldsymbol{TrooperRace}_i \ \ \boldsymbol{DriverRace}_i + \gamma X_i + \varepsilon_i$$

In this model, the coefficient $\boldsymbol{\beta_{rt}}$ represents the marginal difference in search probability experienced in stops where a driver of race *r* is stopped by a trooper of race *t*. Given *R* possible racial classifications of both driver and trooper, the total number of race interactions included in the model is R^2.

To test whether specific driver-trooper race combinations are particularly likely to lead to a search, we perform pairwise comparisons of the combinations. In other words, for each driver-trooper race combination *rt*, we use a chi-squared test to compare the differences in the coefficients $\boldsymbol{\beta_{rt}}$ to the coefficients associated with every other driver-trooper race combination $\boldsymbol{\beta_{r't'}}$. (either r≠r' or t≠t', or both). To the extent that the chi-squared test indicates a significant difference between coefficients $\boldsymbol{\beta_{rt}}$ and $\boldsymbol{\beta_{r't'}}$, we conclude that the relationships identified in Model 4 suggest that highway stops with certain driver-officer race combinations are more likely to lead to a search than others.

---

[87] Fixed effects control for heterogeneity in stops and searches by regional deployment of troopers. For example, there was a higher percentage of Blacks stopped between Exits 1 and 3 of the Turnpike than at exits further north in the patrol area of the Moorestown Station. See, *State v. Pedro Soto*, A734A. 2d 350 (N.J. Super. Ct. Law Div. 1996). Also, troopers assigned to regular patrols make more stops than do their counterparts who work in specialized units. For example, stops may have been made by troopers on special details, such as the "Aggressive Driving Unit" or the "Radar Unit" or the "Tactical Patrol Unit." Evidence presented in Soto based on traffic tickets suggested that when compared to stops by regular patrols, Black motorists were less likely to be stopped by these specialized units. See, *State v. Pedro Soto*, A734A. 2d 350 (N.J. Super. Ct. Law Div. 1996).

[88] Id.

Given R possible racial classifications of both driver and trooper, and R^2 interaction terms included in Model 4, we conduct $\frac{(R-1)(R-1)}{2}$ pairwise comparisons[89]. Assuming that there is less bias in within-race suspect-driver race combinations, we attribute significance in the differences between that estimate and an "off-diagonal" estimate to reflect the marginal influence of race in the request for a search. To the extent that interaction terms are dropped from the model due to collinearity, the number of pairwise comparisons is decreased accordingly. For these tests, we report the ratio of the odds ratios for the logit coefficient of each pair, and the Chi-square significance for the test of the differences in the coefficients.

The comparison of odds ratios allows us to report information both on the direction of the difference as well as its magnitude. So, a ratio of less than 1.0 means that the odds ratio (or unexponentiated coefficient) for the "numerator" driver-officer race pairing lower than the corresponding coefficient for the "denominator" race pairing.

V. Results

A. Descriptive Statistics

Table 1 shows descriptive statistics for the stops. This is a universe of all stops by the NJSP from October 1, 2005 through March 31, 2007. Recall that the final Monitor's Report (Report No. 16) was issued in August, 2007, and included an analysis of data from an overlapping period in 2006-7. Most drivers stopped were White (58.4%). Blacks and Hispanics were smaller minorities.

Table 1 Here

Most officers making stops were White and male, and has been serving in the NJSP for more than six years. Based on data on the race and ethnicity of officers, the NJSP appears to be a racially homogeneous policing institution. More than four stops in five were made by White officers. Black and Hispanic officers made approximately 13% of all stops, and Asian and other ethnicity officers made fewer than three percent of all stops. Note that the officer race measure is not a census of officers; many of these stops could be repeat stops by the same officer. Instead, this is in effect an exposure model, where drivers of various races are "exposed" to the same officer among a group of heterogeneous officers whose population parameters are unknown.

Most (61.3%) of the vehicles stopped were from New Jersey or New York (14.4%). There were few stops made of vehicles registered in Pennsylvania (5.5%). Stops made of vehicles registered in other states along the Interstate 95 corridor were 10.3% of all stops, mostly from Virginia (2.09%) and Maryland (3.23%). Whether this is selection or a weighted population estimate given geographical proximity is hard to say.[90]

---

[89] The division by 2 in computing the number of pairwise comparisons is due to the symmetry of the comparisons. The chi-squared tests examine the equality of coefficients; a test of the null hypothesis: $\beta_{rt}=\beta_{r't'}$ is equivalent to a test of the hypothesis $\beta_{r't'}=\beta_{rt.}$

[90] Mention linkages of NJ and MD in the profiling cases and also in the two Consent Decrees.

Tables 2 and 3 show the distribution of stops and searches by vehicle registration and driver race. Black and Hispanic drivers were searched at far higher rates per stop than were White or Other Race drivers. The orders of magnitude are high – Blacks are searched at nearly three times the rate of White drivers, and Hispanics at more than twice the rate. Other race drivers, predominantly Asians or Native Americans, are searched less often than are White drivers. Vehicle searches by state reflect the distribution of stops, with exception of cases where the vehicle registration is omitted. Out of state plates from the other non-adjacent states are searched in proportion to the frequency with which they are stopped.

Tables 2 and 3 Here

Figure 1 shows the division of the state into trooper stations. We focus on the Moorestown station in portions of this analysis. NJSP troopers assigned to the Moorestown Station patrol the southernmost portion of the Turnpike, from exits 1 to 3. Readers unfamiliar with the state can note that the state's population centers are in the northern portion of the state, from Exit 10 in New Brunswick north to Exit 18 near the intersection with Interstate 80, the highway that leads to the George Washington Bridge that connects to New York City. Traffic on that portion of the Turnpike is predominantly within-state, or within the "tri-state" region of New York, New Jersey and Connecticut.

Insert Figure 1 Here

In the Southern portion of the state, the traffic is less dense as is the population, and vehicles more likely to be transients coming and going from out-of-state. Beyond any demographic mismatch of local drivers in this region with the surrounding population, the focus of stops in this portion of the Turnpike on a transient population of out-of-state vehicles is consistent with DEA guidelines and other New Jersey institutional preferences on the indicia of vehicles potentially trafficking in drugs or other contraband.[91] In fact, Moorestown station stops of vehicles registered in non-adjacent states along the I-95 corridor were 27.8% of all Moorestown stops, compare to fewer than 10% of the total stops. Maryland registrations alone accounted for 10.9% of Moorestown stops, compared to 3.23% statewide.

The frequency of stops statewide and in Moorestown varies over time, but show a steady growth statewide from approximately 37,000 to 50,000 stops over the six calendar quarters that we examined. Stops in Moorestown rose and fell over the six quarters, twice peaking at over 7,800 stops per quarter before falling again in the next quarter. Overall, across the six quarters, Moorestown stops rose from about 6,900 to 7,700.

Figure 2 Here

B. Race and Searches

[91] DEA Guidelines CITE. CITE Veneiro memoranda, acknowledgement of DEA training materials

We estimated a series of logistic regressions to identify the contributions of race of both driver and officer to the probability of being searched. Models were estimated for the statewide set of cases and also for the Moorestown subset.

The models in Table 4 contrast the effects of driver race with the combined effects of driver and trooper race on the probability of a search. We estimate these models with and without fixed effects for NJSP organizational unit to estimate the influences of both location and local culture. Additional models use a global measure of mismatch in officer/driver race. We then include driver race with the mismatch variable to determine if mismatches have effects that are neutral with respect to driver race. The tables show the odds ratio (the exponentiated logit coefficient) for ease of interpretation, together with the standard error for the unexponentiated parameter estimate.

Table 4 Here

Driver race alone predicts the probability of a search. We omitted the White driver category for reference. The results are substantively the same with and without fixed effects for organizational unit, though the coefficients are slightly higher. In models 1 and 2, Black drivers are more than 2.8 times more likely to be searched than White drivers. Latino drivers are more than 2.5 times more likely to be searched than White drivers. Drivers of other races, primarily Asians, are significantly less likely than Whites to be searched. When we add in covariates for the characteristics of the vehicle and the officer, including officer race, these results remain substantively the same. There is a slight decline in the odds ratio for Latino drivers, but it remains significant. There is a slight increase in the odds ratio for Black drivers once we control for officer race and other factors. Stops for reasons other than moving violations are far more likely to lead to searches. The vague category of "other reasons" leads to searches nearly 15 times more than stops for moving violations, suggesting that these stops may in fact be pretexts to conduct a search absent some other codified reason for the stop.

The results for officer race in Table 4 are significant only for Black officers, and the odds ratio is .812. This suggests that Black officers are nearly 20% less likely to conduct a search given a stop, compared to their White counterparts. Other officer races were not significant. One way to interpret this effect is simply that Black officers have a higher threshold of "suspicion" when making a stop and deciding whether to conduct a search, or they may have equal suspicion to White officers but are more temperate in their exercise of discretion.

Models 5 and 6 suggest that a driver-officer race mismatch is less likely to result in a search, and that is consistent even after controlling for driver race. Black and Latino drivers are still more likely to be searched than are White drivers and other race drivers, even when mismatch is added to the models. The odds ratios are consistent with the magnitudes in Models 3 and 4, suggesting that the driver race finding is robust to specifications that take into account the race of the officer.

So far, then, there is fairly strong evidence that Black and Latino drivers are subject to search more often than are White and other race drivers. Table 5 shows results of specific models for each driver race, which begins the process of identifying specific officer-driver race

combinations that might produce differences in search probabilities. The models are specified both with and without fixed effects for organizational unit and calendar quarter. Results are similar though not identical for these alternate specifications.

Table 5 Here

White drivers are less likely to be searched by Black officers compared to White officers, but only without controls for organizational unit. Overall, Black officers are neutral compared to their White counterparts to search drivers, regardless of driver race. The odds ratios are well below 1.0, although most specifications are not significant. Hispanic officers are slightly more likely to search White drivers than other drivers (OR=1.123, $p < .050$) compared to White officers, but that is the only condition where their search probability is significantly different than White officers.

Other race officers (again, mostly Asian) are less likely to search White and Latino drivers, and more likely to search Black drivers. Their OR for searching other race drivers is astonishingly low (.032), suggesting the virtual absence of searches.

All drivers are far more likely to be searched if stopped for the vague category of "other stop reasons," and less likely to be searched if their vehicle is registered in NY compared to New Jersey of Pennsylvania. Finally, troopers with fewer years of service are more likely to conduct searches of all drivers, suggesting that discretion may come with experience and, perhaps, less need for engaging in the types of enforcement that are status-enhancing in the eyes of their professional peers.

The results for stops made by officers assigned to the Moorestown station are nearly identical to stops made elsewhere. Whatever else the Consent Decree may or may not have accomplished, it did produce consistency between this unit and the others in the state. Unfortunately, the pattern of results in Table 6 suggests that Moorestown stops continue the pattern of racial discrimination that was evident before the Consent Decree, and the rest of the NJSP troopers appear now engage in similar patterns of racially selective enforcement.

Table 6 Here

Models 1 and 2 in Table 6 shows that Black and Hispanic drivers, compared to White drivers, are three times more likely to be searched. Again, searches in the "other reason" stop category are many times more likely to take place compared to moving violations stops. The race-specific officer trends suggest Black officers are about 30% less likely to conduct searches in Moorestown compared to White officers, regardless of driver race. Their preference for search is lower across the board, though there are differences by driver race within their negative preferences. Again, we cannot disentangle the reasons for their lower search rate in Moorestown, and it could be that their threshold for suspicion is higher or their taste for discrimination lower. Unlike Black troopers in Moorestown, Latino and other race officers in Moorestown have similar search propensities compared to White troopers.

Table 7 shows tests for specific combinations of officer-driver race combinations. The

statewide set of cases and the Moorestown subsample were analyzed separately. To generate these coefficients, we introduced dummy variables for each driver-officer race combination. We omitted the other race driver/other race officer category as the reference, since it is very small category and drops out of the models. The regressions in Table 7 shows only the odds ratio for each combination, and the standard error for the unexponentiated regression coefficient. All models were estimated with controls for reason for the stop, state of the vehicle's registration, and trooper years of service. Table 7a includes unit fixed effects.

Tables 7a and 7b Here

Comparing the odds ratios in the columns of Table 7a, searches of Black drivers were three to five times more likely to occur compared to searches of White drivers. Officer race did not modify the race-specific search odds. Searches of Latino drivers were also far more likely than searches of White driver, ranging from about twice as likely for Black officers to more than three times more likely for Latino officers.

Patterns in Moorestown suggest a different and more racially distinct set of search practices. White officers were three times more likely to search Black and Latino drivers in Moorestown, but Black officers in the Moorestown station were no more likely to search Black or Latino drivers than they were to search White drivers. Latino officers were more likely to search Latino drivers and Black drivers, though the large odds ratio for that comparison (2.421) was not significant. Other race officers in Moorestown were strongly inclined to search Black drivers, but not drivers of other races. The patterns in Table 7b suggest that conditions in Moorestown have changed little in the decade between the Soto research in 1996 and the closing months of federal oversight in 2006-7.

C. Pairwise Comparisons and Differences

We conducted a series of pairwise comparisons to identify in more detail differences in officer-driver race interactions in search probabilities. We modified the logistic regression models in Table 7 to include specific pairwises test of each driver-officer race combination, and conducted a set of chi-square tests for the differences in parameter estimates for each combination. To help in interpretation, we report two statistics from these tests. First is simply the significance of the bivariate test, or p(Chi-square). Next, the numerical values in the tables are the ratios of the odds ratios for each pair to provide information on the magnitude and direction of the differences. A ratio greater than 1.0 indicates that the ratio for the pair in the row is greater than the odds ratio for the pair in the column. So, whereas Table 7 reports the odds ratio of each combination, Table 8 reports the ratio of the separate odds ratios for each of the two driver-officer race combinations. This statistic is similar to a comparison of excess risk ratios in epidemiology, and allows the comparison of two race pairings' risk to each other, rather than to a previously designated reference group.[92]

[92] Kenneth Rothman, Sander Greenland, and Timothy L. Lash, Modern Epidemiology (2008); See, also, Sana Loue, Sana Loue, *Epidemiological Causation in the Legal Context: Substance and Procedures*, in (J. Gastwirth, ed.) Statistical Science in the Courtroom 254 (2000); A. Laupacis, D.L. Sackett, and RS Roberts, An Assessment of Clinically Useful Measures of the Consequences of Treatment. 318 *N Engl*

Excess risk measures the association between a specified risk factor (in this case, driver/officer race combination) and a specified outcome (the probability of being searched). While epidemiologists generally report differences, here we report the ratio of the two odds ratios. It is the ratio between two proportions—in epidemiology it's typically defined to be the difference between the proportion of subjects in a population with a particular disease who were exposed to a specified risk factor (P(D|E)) and the proportion of subjects with that same disease who were not exposed. Here, we substitute ratio for difference.

Table 8 Here

Ratios greater than 2.0 are highlighted in the table in bold, and consider these to be meaningful differences. Many other differences achieve statistical significance, but in a large set of cases, this may simply be differences without substantive meaning. The 2.0 odds ratio threshold reflects a standard of causation in tort law where courts have recognized that risk can at some threshold translate into causal inference.[93] We also highlight (in italics) ratios of less than .5, which suggest a meaningful difference in the opposite direction (i.e., where the odds ratio of the column pair more than twice exceeds the odds ratio in the row pair). Because the chi-squared tests reflected in Table 8 are tests for equality, and are therefore symmetric, Table 8 has a similar symmetry, where odds ratios of 2.0 or greater have "reflecting" odds ratios of .5 or less when the comparison is reversed.

The results in Table 8 confirm what we observed in the previous tables. Nearly every pair that comparing a Black or Latino driver to a White driver has a ratio of odds ratios greater than 2.0. These patterns hold regardless of the officer race. Similarly, nearly every comparison with a White driver suggests that they are significantly less likely to be searched regardless of the race of the officer. Remember that these are pairwise comparisons, so that each one of these compares the D/O combination to every other D/O combination.

MORE

J Med 1728-33 (1988); Bruce Levin and Michael O. Finkelstein, Statistics for Lawyers (2002); Lee Epstein & Gary King, The Rules of Inference, 69 *Univ. Chic. L. Rev*. 1 (2002).

[93] See, *Landrigan v. Celotex*, 127 N.J. 404; 605 A.2d 1079 (1992) (stating that "Without expressly adopting a specific standard, the court cited with approval several cases that adopted a requirement that an epidemiological study show a relative risk in excess of 2.0 to prove that causation in a specific individual was more probable than not. The significance of a relative risk greater than 2.0 representing a true causal relationship is that the ratio evidences an attributable risk of more than fifty percent, which means that more than half of the cases of the studied disease in a comparable population exposed to the substance are attributable to that exposure. This finding could support an inference that the exposure was the probable cause of the disease in a specific member of the exposed population.") See, also, Sana Loue, supra note _; Levin and Finkelstein, supra note __; Carl Cranor, *Regulating Toxic Substances: A Philosophy of Science and the Law* 1 (1993).

D. Summary

We used a series of regressions to estimate the extent to which there was differential treatment of drivers of different races by the New Jersey State Police in the closing months of a lengthy period of federal oversight and monitoring of their activities. Our metric of differential treatment was the request for a search conditional on being stopped, an indicator of suspicion that would animate the second stage inquiry of a vehicle search. We exploit the heterogeneity of officer-driver race combinations to identify the magnitude of any observed differences.

We developed consistent evidence of racially selective enforcement, with Blacks and Latinos more likely to be searched compared to White drivers. Officer race effects vary according to the specification, but in general, White officers have a greater propensity to stop non-White drivers. We also see evidence that Black officers are more likely to cabin discretion and/or employ a different or higher threshold of suspicion before conducting a search. In general, the coefficients for Black officers in predicting search were lower in nearly all conditions than were the regression parameters for White, Latino or Other Race officers. These results were robust to several different specifications, including tests that included factors related to the vehicle and controlling for the officer's length of service (and therefore, maturity and experience).

We verified these results using two forms of pairwise tests: one test that examined the likelihood of a search given permutations of officer-driver race combinations, and a second that tested each of these permutations against each other. In both instances, we observe significantly higher search probabilities for Black and Latino drivers compared to White and Other Race drivers, regardless of the race of the officer. Conversely, we found significantly lower rates of search requests for White drivers, regardless of officer race. Assuming that same-race officer-driver combinations were the least biased condition, every other condition testing the marginal effect of driver race in predicting a search. So, while at the means, White drivers are more inclined to request searches than Black officers, these effects disappear at the margins when we control for officer-driver race combinations and directly compete them to determine relative risk, or what is known in epidemiology as excess risk. These results also were robust to several covariates.

Of course, these differentials may reflect unobservables attendant to either the stop itself (e.g., whether the driver is an "egregious" speeder who exceed the speed limit by more than 15 miles per hour) or to observables (but not quantifiables) during the stop, such as demeanor, condition of the automobile, passengers, and other factors that may be correlated with race. Inchoate suspicion may be hidden or signaled but not defined in a way that trooper can categorize, and we see this at work in the very large effects we observe when we control for a dummy variable about whether the basis for the stop was "other." To see how this distributes bu driver race, we estimated this condition against a baseline of stops for moving violations. The effects are consistently large and significant: "other reason" is more likely to be invoked with Black drivers (OR=1.44) and Latino (OR=1.64) compared to White drivers. There are no significant differences in the use of this category by officer race. These differentials suggest the importance of indicia of suspicion and discretion that are not "articulable" and thus pose some constitutional dilemmas in terms of contemporary expressions of "reasonable suspicion" that

govern searches.

Since searches are conditional on stops, we also should consider as a conditioning process the results from the Advisory Committee's deliberations on new evidence of racial selection of drivers for stops.[94] When we couple robust evidence of race differences in search requests with the LK results on race differentials in the selection of drivers for stops, we find a compounding of bias in stops that defies the terms and conditions of the Consent Decree. We turn next to the institutional and structural architecture of the Consent Decree, and place it in context of the larger set of Consent Decrees and MOAs, to explain the persistence of racially selective enforcement post-*Soto*.

VI. The *Soto* Consent Decree and the Institutional Design of Oversight

A. The challenge in any effort for institutional reform of policing is the design of procedures and policies that can effectively manage police discretion. This requires measures that minimize both "rotten apples" and replaces the institutional environment that produced them. (Armacost).

B. But the design of Soto and other consent decrees suggests that even if we are better at managing discretion, there is a residual component of discretion that is beyond the reach of training, oversight, and professionalization. The challenge for both leadership, monitors, and social scientists is to identify the moving parts of racially biased policing that seem to evade institutional reforms.

C. So, given the brief history of the impacts of consent decrees, the Soto experience becomes an important case study on legal regulation of policing. Soto was an exceptionally well crafted design --- DETAILS. Any limitations of Soto are likely to pop up in future consent decrees that repeat this design.

So, we ask why did discriminatory patterns persist in the wake of the Soto, and are there flaws in the design of the Consent Decree that could achieve positives in all but the critical benchmark of racially disparate treatment?

- Design of the Consent Decree itself
- Weaknesses in the monitoring function – looking in the wrong places, weak analyses
- Less than full disclosure in the reporting of data, perhaps problems were always present but undetected because the monitor lacked the chops to do so.
- Sociology of monitoring, narrowing of cognitive frames for oversight
- Networks of monitors undermines "independence" – defeats the broader goals of democratic regulation that co-exists with legal regulation

[94] LK study, supra note __.

D. **Why the failure to benchmark police accountability**? Why not a disparate treatment test as a routinized dimension of monitoring? What are we afraid of?

E. Prospects for legal regulation are dim. While we move toward an evidence-base in a broad range of social policies, from medicine to criminal justice, the reluctance to impose harder benchmarks in this context is puzzling. We instead look to procedure, training and professionalization as a remedy to factors that may operate outside of those structures, and may result from cognitive or unconscious processes that are unlikely to be reshaped by focusing on reifying professional standards of conduct

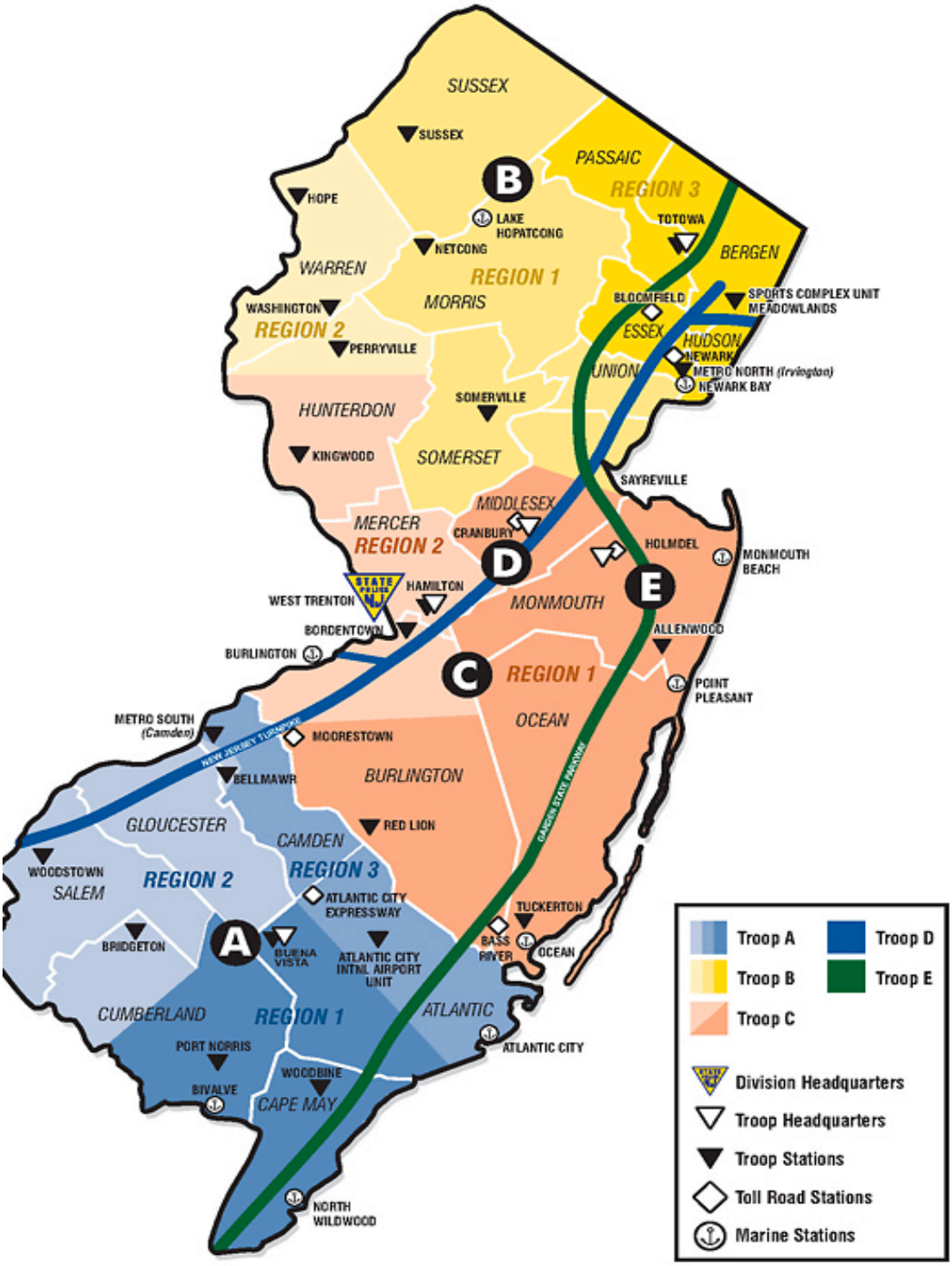
F. Seeking an alternative – democratic regulation, scientific as well as professional oversight, use of social framework models, and taking seriously the notion of democratic accountability and citizen participation. See Fung on Chiago CAPS, other perspectives on CAPS (Skogan), San Diego (Cordner), Boston Ten Point Coalition, other examples. Move beyond policing to other domains, including environmental regulation and education. See Dorf and Sabel (2001), Sabel and Simon (2009)

VI. Conclusion

Table 1. Descriptive Statistics (Mean, SD)

	Drivers		***Troopers***	
Race (%)				
White	58.40	(49.28)	84.55	(36.14)
Black	19.97	(39.97)	4.43	(20.57)
Hispanic	13.21	(33.87)	8.56	(27.97)
Asian	4.50	(20.72)	1.12	(10.51)
American Indian	0.68	(8.24)	1.26	(11.16)
Pacific Islander	2.45	(15.44)	0	0.00
Unknown	0.79	(8.86)	0	0.00
Gender (%)				
Male	77.14	(41.99)	97.18	(16.57)
Female	22.86		2.82	
Age	35.29	(12.65)		
Years of Service			6.28	(4.58)
Vehicle Registration				
New Jersey	61.28	(48.71)		
New York	14.42	(35.13)		
Pennsylvania	5.48	(22.77)		
Other Out-of-State	15.63	(36.31)		
Missing	3.19	(17.57)		
Reason for Stop				
Moving Violation	89.02	(31.27)		
Non-Moving Violation	9.68	(29.57)		
Other	1.30	(11.33)		

Figure 1. New Jersey State Police Patrol Sectors



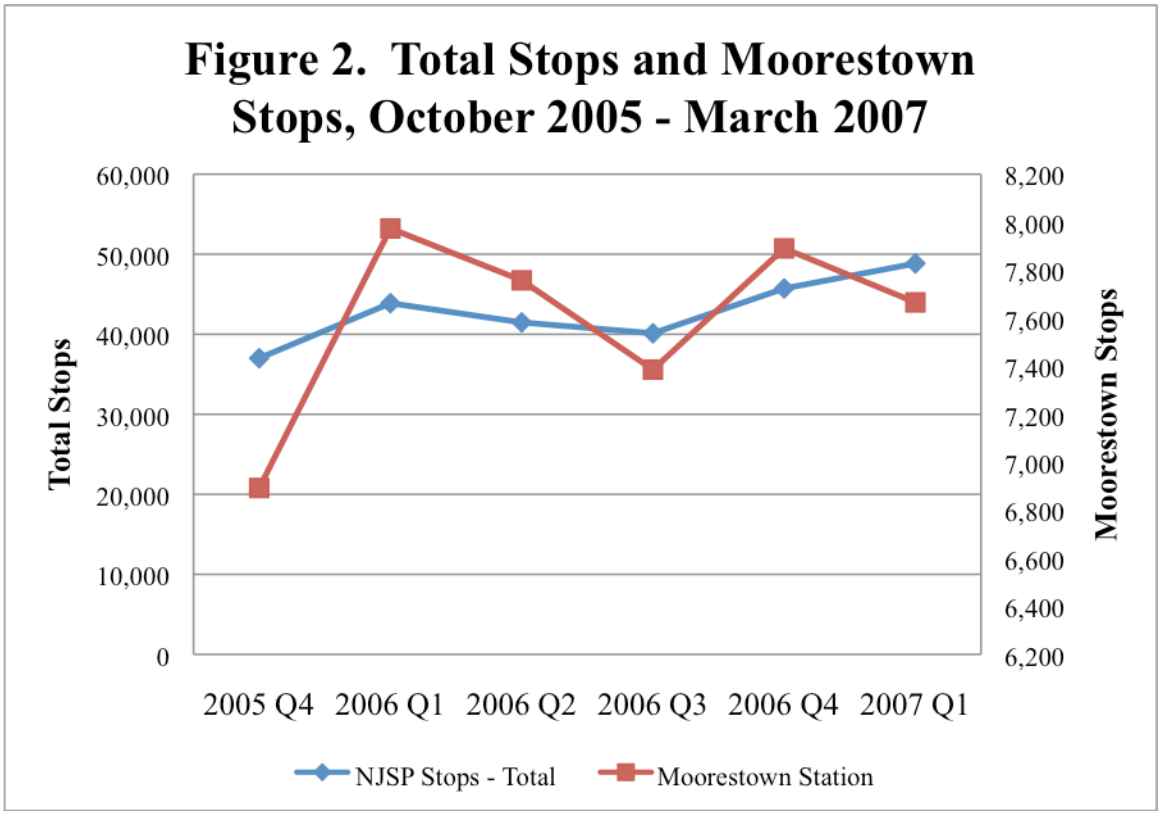


Figure 2. Total Stops and Moorestown Stops, October 2005 - March 2007
Total Stops
60,000
50,000
40,000
30,000
20,000
10,000
0
8,200
8,000
7,800
7,600
7,400
7,200
7,000
6,800
6,600
6,400
6,200
Moorestown Stops
2005 Q4
2006 Q1
2006 Q2
2006 Q3
2006 Q4
2007 Q1
NJSP Stops - Total
Moorestown Station

Table 2. Stops and Searches by Driver Race, New Jersey State Police, October 2005-March 2007

		Driver Race			
	Total	*White*	*Black*	*Hispanic*	*Other*
N of stops	257,059	150,131	51,325	33,970	21,633
% of All Stops	100.0	58.4	20.0	13.2	8.4
N of searches	10,890	4,400	3,829	2,235	426
% Searched	4.2	2.9	7.5	6.6	2.0

Table 3. Stops and Searches by Vehicle Registration, New Jersey State Police, October 2005- March 2007

		Vehicle Registration				
	Total	*New Jersey*	*New York*	*Pennsylvania*	*Other*	*Missing*
N of stops	257,059	157,531	37,069	14,096	40,168	8,195
% of All Stops	100	61.3	14.4	5.5	15.6	3.2
N of searches	10,890	6,537	823	644	1,435	1,451
% Searched	100	60.0	7.6	5.9	13.2	13.3

Table 4. Logistic Regression of Search Probability by Driver Race, Trooper Characteristics, Stop Characteristics, and Vehicle Characteristics (Odds Ratio, SE)

Model	(1)	(2)	(3)	(4)	(5)	(6)
Predictors	Driver Race Only	Driver Race Only	Driver and Trooper Race	Driver and Trooper Race	Race Mismatch	Driver Race and Race Mismatch
Constant	.030	.032	.042	.035	.045	.042
	(0.015)**	(0.031)**	(0.024)**	(0.037)**	(0.024)**	(0.024)**
Driver - Black	2.670	2.832	2.664	2.754		2.743
	(0.023)**	(0.024)**	(0.024)**	(0.025)**		(0.036)**
Driver - Hispanic	2.333	2.555	2.179	2.366		2.239
	(0.027)**	(0.028)**	(0.028)**	(0.029)**		(0.038)**
Driver - Other Race	.666	.698	.660	.689		.680
	(0.051)**	(0.052)**	(0.053)**	(0.053)**		(0.060)**
Trooper Years of Service			.921	.949	.919	.921
			(0.003)**	(0.003)**	(0.003)**	(0.003)**
Stop - Non-moving Violation			2.535	2.743	2.593	2.537
			(0.026)**	(0.027)**	(0.026)**	(0.026)**
Stop - Other Reason			17.305	15.394	17.567	17.305
			(0.039)**	(0.041)**	(0.039)**	(0.039)**
Vehicle - New York			.529	.536	.524	.528
			(0.038)**	(0.038)**	(0.038)**	(0.038)**
Vehicle - Pennsylvania			1.042	1.097	1.058	1.041
			(0.043)	(0.044)*	-0.043	-0.043
Vehicle - Other Non-NJ			.687	.764	.731	.687
			(0.031)**	(0.032)**	(0.030)**	(0.030)**
Driver-Officer Race Mismatch					1.732	0.962
					(0.021)**	(0.034)
Trooper - Hispanic			.983	1.043		
			(0.035)	(0.037)		
Trooper - Other			1.037	1.009		
			(0.062)	(0.063)		
Trooper - Black			.810	.812		
			(0.053)**	(0.054)**		
Observations	257,059	255,824	256,837	255,604	256,837	256,837
Unit FE?	No	Yes	No	Yes	No	No

Significance: * = p < .05, ** = p < .01

Table 5. Logistic Regression of Race-Specific Search Probability by Trooper Characteristics, Stop Characteristics, and Vehicle Characteristics (Odds Ratio, SE)

Model	(1)	(2)	(3)	(4)	(5)	(6)	(5)	(6)
Predictors	White Drivers		Black Drivers		Hispanic Drivers		Other Race/Ethnicity Drivers	
Constant	0.039	0.033	0.118	0.089	0.092	0.097	0.035	0.770
	(0.032)**	(0.058)**	(0.037)**	(0.061)**	(0.048)**	(0.068)**	(0.102)**	(0.196)
Trooper Years of Service	0.931	0.964	0.916	0.933	0.916	0.950	0.906	0.928
	(0.004)**	(0.005)**	(0.005)**	(0.006)**	(0.007)**	(0.007)**	(0.015)**	(0.016)**
Stop - Non-moving Violation	2.550	2.924	3.093	3.108	1.863	2.052	2.010	2.335
	(0.043)**	(0.044)**	(0.043)**	(0.044)**	(0.058)**	(0.060)**	(0.138)**	(0.145)**
Stop - Other Reason	24.288	21.200	11.336	10.044	13.749	12.730	19.279	18.728
	(0.056)**	(0.059)**	(0.075)**	(0.078)**	(0.087)**	(0.091)**	(0.160)**	(0.169)**
Vehicle - New York	0.538	0.564	0.432	0.437	0.708	0.679	0.531	0.460
	(0.062)**	(0.062)**	(0.067)**	(0.068)**	(0.079)**	(0.080)**	(0.152)**	(0.154)**
Vehicle - Pennsylvania	1.012	1.096	0.875	0.940	1.573	1.499	0.400	0.407
	(0.071)	(0.073)	(0.076)	(0.078)	(0.081)**	(0.083)**	(0.297)**	(0.301)**
Vehicle - Other Non-NJ	0.611	0.744	0.624	0.672	1.017	1.023	0.494	0.526
	(0.058)**	(0.061)**	(0.046)**	(0.050)**	(0.062)	(0.066)	(0.158)**	(0.165)**
Trooper - Hispanic	1.015	1.123	0.942	1.007	1.039	1.025	0.792	1.266
	(0.056)	(0.058)*	(0.061)	(0.063)	(0.074)	(0.076)	(0.190)	(0.287)
Trooper - Other	0.755	0.782	1.379	1.302	0.882	0.829	1.284	0.032
	(0.117)*	(0.119)*	(0.092)**	(0.093)**	(0.140)	(0.141)	(0.282)	(0.156)**
Trooper - Black	0.810	0.874	0.856	0.865	0.811	0.738	0.567	0.525
	(0.088)*	(0.089)	(0.085)	(0.086)	(0.114)	(0.115)**	(0.283)*	(0.287)*
Observations	150,018	145,642	51,259	49,755	33,951	32,404	21,609	17,532
Unit FE?	No	Yes	No	Yes	No	Yes	No	Yes

Table 6. Logistic Regression of Search Probability by Driver Race, Trooper Characteristics, Stop Characteristics, and Vehicle Characteristics: Moorestown Unit (Odds Ratio, SE)

Model	(1)	(2)	(3)	(4)	(5)	(6)
Predictors	Driver Race Only	Driver Race with Controls	White Drivers	Black Drivers	Hispanic Drivers	Other Race Drivers
Constant	.024	.039	.031	.139	.117	.049
	(0.046)**	(0.070)**	(0.107)**	(0.086)**	(0.138)**	(0.307)**
Driver - Black	2.986	3.149				
	(0.058)**	(0.059)**				
Driver - Hispanic	3.174	3.155				
	(0.071)**	(0.073)**				
Driver - Other Race	.689	.736				
	(0.132)**	(0.134)*				
Trooper Years of Service		.901	.945	.884	.901	.796
		(0.010)**	(0.017)**	(0.015)**	(0.021)**	(0.066)**
Stop - Non-moving Violation		2.392	2.326	2.344	2.575	2.612
		(0.075)**	(0.136)**	(0.113)**	(0.166)**	(0.426)*
Stop - Other Reason		18.412	32.819	9.689	25.053	23.220
		(0.103)**	(0.179)**	(0.159)**	(0.224)**	(0.439)**
Vehicle - New York		.491	.449	.436	.707	.464
		(0.087)**	(0.183)**	(0.125)**	(0.183)	(0.386)*
Vehicle - Pennsylvania		.892	1.213	.843	.764	.424
		(0.092)	(0.156)	(0.140)	(0.209)	(0.538)
Vehicle - Other Non-NJ		.709	.676	.711	.770	.434
		(0.059)**	(0.118)**	(0.081)**	-0.135	(0.324)**
Trooper - Black		.597	.754	.490	.235	1.885
		(0.135)**	(0.234)	(0.206)**	(0.514)**	(0.397)
Trooper - Hispanic		.912	.987	.844	1.151	
		(0.107)	(0.195)	(0.158)	(0.221)	
Trooper - Other		1.174	.808	1.548	.787	1.618
		(0.095)	(0.219)	(0.123)**	(0.242)	(0.479)
Observations	42,837	42,837	20,180	13,421	5,236	3,850

Significance: * = p < .05, ** = p < .01

Table 7a. Logistic Regression of Search by Driver/Officer Race and Ethnicity Combinations (Odds Ratio, SE)

		Driver Race			
		White	***Black***	***Latino***	***Other***
Officer Race	***White***	1.346	3.553 ***	2.999 ***	0.915
		(.377)	(.996)	(.842)	(.261)
	Black	1.063	2.911 ***	2.276 **	0.500
		(.311)	(.849)	(.686)	(.196)
	Latino	1.440	3.473 **	3.244 ***	0.745
		(.410)	(.993)	(.936)	(.248)
	Other	0.927	4.735 ***	2.462 **	--
		(.280)	(1.390)	(.769)	

Significance: a = p < .000, b = p< .01, c = p < .05

Estimates controlled for vehicle state license plate, trooper years of service, reason for stop and unit fixed effects. Model for Other Race Driver/Other Race Officer omitted, N=24

N = 256,501

Log Likelihood = - 39142

Pseudo R^2 = .122

Table 7b. Logistic Regression of Search by Driver/Officer Race and Ethnicity Combinations, Moorestown Only (Odds Ratio, SE)

		Driver Race			
		White	***Black***	***Latino***	***Other***
Officer Race	***White***	0.993	3.121 *	3.118 *	0.734
		(.452)	(1.419)	(1.457)	(.367)
	Black	0.716	1.546	0.769	1.469
		(.363)	(.768)	(.522)	(.829)
	Latino	1.440	2.421	3.406 *	--
		(.410)	(1.159)	(1.698)	
	Other	0.927	5.049 ***	2.558	--
		(.280)	(2.356)	(1.291)	

Significance: a = p < .000, b = p< .01, c = p < .05

Estimates controlled for vehicle state license plate, trooper years of service, and reason for stop. Model for Other Race Driver/Other Race Officer and Other Race Driver/Latino Officer omitted, no variation

N = 45,409

Log Likelihood = - 7986

Pseudo R^2 = .116

Table 8. Pairwise Tests of Differences in Parameter Estimates (Ratio of Odds Ratios, p(Chi-square))

	WD/WO	***WD / BO***	***WD / LO***	***WD / OO***	***BD / WO***	***BD / BO***	***BD / LO***	***BD /OO***	***LD / WO***	***LD / BO***	***LD / LO***	***LD / OO***
WD / WO		1.267 b	NS	1.464 b	*.379 a*	*.464 a*	*.388 a*	*.287 a*	*.449 a*	.592 a	*.415 a*	.547 a
WD / BO	.789 b		.739 b	NS	*.299 a*	*.366 a*	*.306 a*	*.226 a*	*.355 a*	*.468 a*	*.328 a*	*.432 a*
WD / LO	NS	1.353 b		1.565 a	*.405 a*	*.496 a*	*.415 a*	*.306 a*	*.480 a*	.633 b	*.444 a*	*.585 b*
WD / OO	.683 b	NS	.639 a		*.259 a*	*.317 b*	*.265 a*	*.196 a*	*.307 a*	*.404 a*	*.306 a*	*.374 a*
BD / WO	**2.638 a**	**3.342 a**	**2.469 a**	**3.863 a**		1.223 c	NS	.757 b	1.186 a	1.563 b	NS	1.444 b
BD / BO	**2.155 a**	**2.732 a**	**2.018 a**	**3.159 b**	.818 c		NS	.619 b	NS	NS	NS	NS
BD / LO	**2.577 a**	**3.268 a**	**2.411 a**	**3.774 a**	NS	NS		.740 b	1.157 c	1.527 b	NS	1.410 c
BD /OO	**3.484 a**	**4.425 a**	**3.263 a**	**5.106 a**	1.321 a	1.616 b	1.353 b		1.567 a	**2.065** a	1.448 a	1.908 b
LD / WO	**2.227 a**	**2.817 a**	**2.082 a**	**3.257 a**	.843 a	NS	.864 c	.638 a		1.318 c	NS	NS
LD / BO	1.689 a	**2.317** a	1.580 b	**2.475** a	.640 b	NS	.655 b	.484 a	.759 c		.701 b	NS
LD / LO	**2.410 a**	**3.049 a**	**2.254 a**	**3.268** a	NS	NS	NS	.691 a	NS	1.427 b		NS
LD / OO	1.828 a	**2.315** a	1.709 b	**2.674** a	.693 b	NS	.709 c	.524 b	NS	NS	NS	

Table entries represent the ratio of the "row" odds ratio to that of the "column" odds ratio. "NS" entries represent differences that are statistically insignificant.
Significance: a = p < .000, b = p < .01, c = p < .05

Notation:

LD: Latino Driver	LO: Latino Officer
BD: Black Driver	BO: Black Officer
WD: White Driver	WO: White Officer
OD: Other Race Driver	OO: Other Race Officer

Appendix A. Monitor Function Key

Key:
§ 14141 Suit
§ 14141 Investigation – consent decree before lawsuit
§ 1983 Suit
State AG or Municipal Suit

A. Private Monitoring by Security Firms
Detroit – 2003 – Sheryl Robinson Wood – removed by judge for misconduct in 2009, replaced by Robert Warshaw (see below)
Los Angeles – 2001 – Michael Cherkasky (NYCDA, Kroll), deputy Jeff Schlanger (NYC DA, private practice, private investigation firm purchased by Kroll) – extended in 2006, final report in 2009 with unfulfilled issues (including racial profiling) to be taken on by Office of the Inspector General
Austin – 2007 – Kroll hired to audit an IAB report of a shooting. Team, led by Jeff Schlanger, found that the police officer had failed to follow protocol before the shooting, but did not rule on the shooting itself, seconding the findings of the IAB investigation.
Pennsylvania – 2003 – Investigation at request of Governor Rendell to determine department's compliance with new sexual harassment policy. Four reports issued over one year by Sheryl Robinson Wood, concluding that the department was in compliance

B. DOJ Monitoring its own agreements
Buffalo – 2002 – no evidence of DOJ followup and no notice that monitoring has ended
Cleveland – 2004 – One-year period of DOJ full unrestricted access – nothing published or granted afterwards
Montgomery Co. MD – 2000 – Agreement required annual reporting on racial breakdown of traffic stops. After first year, no reports made public. No action by DOJ.
Mt. Prospect, IL – 2003 – MOU to end investigation, department must file a report with DOJ every six months. No public documentation of such reports available.
Villa Rica, GA – 2003 – DOJ investigated and entered in an MOU stating that Villa Rica needed to assign an auditor to oversee compliance with racial profiling policy. No info on DOJ website, PARC, or Civil Rights Litigation Clearinghouse suggesting that a monitor was ever selected.

C. Individual Monitors – Private Practice Attorneys
Cincinnati – 2002 Saul Green (Former US Atty, in private practice in 2002, now Deputy Mayor of Detroit) – Joseph Brann is deputy monitor
District of Columbia – 1999 (by request of department) – Michael Bromwich (SDNY, OIC, Inspector General of DOJ, Fried Frank) – in 2008 Bromwich stated DC was in compliance with 80% of the 126 substantive provisions.
US Virgin Islands – 2009 – Michael Bromwich appointed and issued one very critical report in May 2010. It is not clear who is monitoring the USVI now that Bromwich has

taken a position in the Obama administration.

D. Private Monitoring Firms
Pittsburgh – 1997 – **Police Management Resources** (James Ginger). Nineteen quarterly reports issued, through 2002, when the decree was lifted.
New Jersey – 1999 – **Police Management Resources** (James Ginger). Final report in April 2009 states department has become "self-monitoring" and consent decree lifted in September 2009.
Oakland – 2003 – Four-person team (two former police officials from other cities, two civil rights attorneys) appointed to oversee 52 "tasks" in settlement. Decree extended for failure to comply in 2007. Team issued 14 reports over five years. Team replaced by Robert Warshak and **Police Performance Solutions**
Prince George's County – 2004 – Complaints of excessive force by Canine Unit. Initially DOJ monitored, issuing joint reports. Contracted with The Alexander Group, a subdivision of MPRI, part of L3 (international military contractor), which issued reports until January 15, 2009, when consent decree was lifted.
Riverside, CA – 2001 – State AG suit, stipulated that state AG could hire a consultant. State AG hired Joseph Brann, the former police chief of Santa Ana and Saul Green's deputy in Cincinnati. Brann had his own private consulting firm at the time, and now runs Public Strategies Group, a government consulting firm.
Steubenville, OH – 1997 – DOJ chose Charles Reynolds. At the time, he did not appear to work at a consulting firm, though he later joined **Police Management Resources** and then co-founded **Police Performance Solutions**. No public reports by this auditor are available.
Walkill, NY – 2001 – Suit by NY State OAG requiring implementation of screening and performance monitoring. OAG and Walkill agreed on Dean Esserman (NYU law, NYC ADA, General Counsel to the Transit Police for Bratton, Assistant Chief for the New Haven Police, Chief of Stamford Police). Since being appointed, he became chief of Providence, RI. He issued four annual reports, asking for six additional months of monitoring before stating that the department was in compliance in January 2006.

Key figures in the Network of Monitors

Robert Warshaw—former Chief of Police in Rochester, former Associate Director of Office of National Drug Control Policy. Senior Associate of **Simeone Associates**, one of three partners in **Police Performance Solutions.**

Charles Reynolds—former Chief of Police of five towns, including Lebanon and Dover NH. He served as the Independent Auditor for Steubenville, OH decree. His bio states that he served on monitoring team in New Jersey and Prince Georges's County, suggesting that he worked at **Police Management Resources**. He is one of three partners in **Police Performance Solutions.**

Joseph Wolfinger—former Assistant Director of Training for the FBI, one of three partners in **Police Performance Solutions.**

Kroll Executives – Schlanger, Cherkasksy

James Ginger – Former Deputy Director of the Police Foundation, founders of **Police Management Resources**.